

**MENU****SEARCH****INDEX****DETAIL****JAPANESE****LEGAL  
STATUS**

1 / 1

**PATENT ABSTRACTS OF JAPAN**

(11)Publication number : 2004-145448

(43)Date of publication of application : 20.05.2004

(51)Int.Cl.

G06T 17/40

G06T 1/00

G06T 3/00

G06T 7/60

G06T 15/70

(21)Application number : 2002-307170

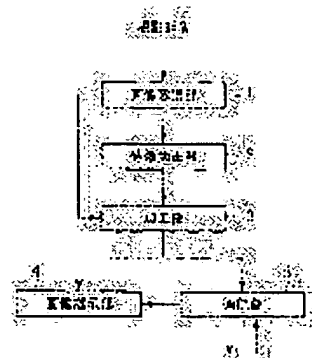
(71)Applicant : TOSHIBA CORP

(22)Date of filing : 22.10.2002

(72)Inventor : MIHARA NORIO  
NUMAZAKI SHUNICHI  
HARASHIMA TAKAHIRO  
KISHIKAWA KUNIHISA  
DOI MIWAKO**(54) TERMINAL DEVICE, SERVER DEVICE, AND IMAGE PROCESSING METHOD****(57)Abstract:**

**PROBLEM TO BE SOLVED:** To provide a terminal device capable of processing images, such as for synthesizing an advertising image or a desired image, according to the three-dimensional features of a plurality of objects in the images(including videos and still pictures), such as their three-dimensional configurations and three-dimensional relationship between their positions.

**SOLUTION:** A first image containing a plurality of objects is obtained and depth information corresponding to a plurality of areas in the first image composed of the plurality of areas made up of one or a plurality of pixels is obtained. Based on the first image and the depth information, at least the three-dimensional features of the plurality of objects and the three-dimensional relationship between the positions of the plurality of objects are extracted as feature information and the first image is processed according to this feature information. As a result, images can be processed to synthesize the advertising image or an image of a desired virtual object based on the three-dimensional features of the plurality of objects in the images (including video and still pictures), such as their three-dimensional configurations and the relationship between their positions.



1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.\*\*\*\* shows the word which can not be translated.

3.In the drawings, any words are not translated.

---

## CLAIMS

---

[Claim(s)]

[Claim 1]

It is a terminal unit which is one of two or more terminal units which can communicate mutually, The 1st acquisition means that acquires the 1st picture containing two or more candidates for photography,

The 2nd acquisition means that said 1st picture comprises two or more fields which consist of 1 or two or more pixels, and acquires depth information corresponding to each of two or more of these fields,

An extraction means to extract three-dimensional physical relationship three-dimensional shape for [ said / two or more ] photography, and for [ said / two or more ] photography as characteristic information at least based on said 1st picture and said depth information,

A creating means which processes said 1st picture and generates the 2nd picture based on characteristic information extracted by this extraction means,

A providing terminal unit.

[Claim 2]

Said creating means is what generates said 2nd picture by combining an objective picture in said 1st picture, Three-dimensional physical relationship of for concerned an object and for [ said / two or more ] photography when a position of the object concerned is provided in a depth direction in the 1st picture concerned at that time, The terminal unit according to claim 1 characterized by combining a picture of the object concerned in the 1st picture concerned based on at least one of three-dimensional shape for [ said / two or more ] photography.

[Claim 3]

By combining an objective picture in said 1st picture, said creating means generates said 2nd picture, and in that case, it controls it so that a motion of said object suits three-dimensional shape for [ said ] photography, The terminal unit according to claim 1 combining a picture of the object concerned in the 1st picture concerned.

[Claim 4]

Said extraction means extracts a position of a depth direction where said candidate for photography in said 1st picture exists, and a background region of the back for [ concerned ] photography,

The terminal unit according to claim 1, wherein said creating means generates said 2nd picture, and it combines a picture of the object concerned in the 1st picture concerned by combining an objective picture in said 1st picture in that case so that the object concerned may exist between said background regions for photography concerned.

[Claim 5]

Said extraction means extracts the 2nd position of a depth direction where the 2nd candidate for photography that is one of the 1st position of a depth direction where the 1st candidate for photography that is one of for [ in said 1st picture / said / two or more ] photography exists, and everything [ the ] but inside for [ in the back 1st for / concerned / photography / said / two or

more ] photography exists,

The terminal unit according to claim 1, wherein said creating means generates said 2nd picture, and it combines a picture of the object concerned in the 1st picture concerned by combining an objective picture in said 1st picture in that case so that the object concerned may exist between the 1st position concerned and the 2nd position.

[Claim 6]

Said extraction means extracts the 3rd position as a three-dimensional position in which said candidate for photography in said 1st picture exists,

When said creating means generates said 2nd picture by combining an objective picture in said 1st picture and the object concerned results in said 3rd position based on a motion of the object concerned and a motion for [ concerned ] photography at that time, The terminal unit according to claim 1 judging that the object concerned collided with the candidate for photography concerned, controlling so that a motion and expression of the object concerned correspond to a collision, and combining a picture of the object concerned in the 1st picture concerned.

[Claim 7]

The terminal unit according to claim 6 when it judges [ said creating means ] that said object collided with said candidate for photography, wherein it compounds effect expression of a collision in said 1st picture.

[Claim 8]

Said extraction means extracts the 3rd position as a three-dimensional position in which said candidate for photography in said 1st picture exists,

By combining an objective picture in said 1st picture, said creating means generates said 2nd picture, and in that case Said 3rd position for [ concerned ] photography, The terminal unit according to claim 1 controlling a motion and expression of the object concerned and combining a picture of the object concerned in the 1st picture concerned based on the 4th position that is a three-dimensional position of the object concerned in said 1st picture.

[Claim 9]

Said extraction means extracts each picture for [ said / two or more ] photography, and a picture of a background region after them from said 1st picture,

The terminal unit according to claim 1, wherein said creating means generates said 2nd picture by processing at least one of a picture of said background region, and each pictures for [ said / two or more ] photography.

[Claim 10]

Said 2nd acquisition means,

A means to acquire the 3rd picture that is equivalent to said 1st picture and includes depth information in each pixel value,

A means to extract depth information corresponding to each field of said 3rd picture to said 1st picture,

It provides,

The terminal unit according to claim 1, wherein said extraction means extracts three-dimensional physical relationship three-dimensional shape for [ said ] photography, and for [ said / two or more ] photography from said depth information corresponding to each field of said 1st picture and the 1st picture concerned at least.

[Claim 11]

Said 2nd acquisition means,

A luminescent means which emits light towards said candidate for photography,

The 1st light-receiving means that receives the 1st light volume containing catoptric light from the candidate for photography concerned of light made this luminescent means emit light and irradiated said candidate for photography,

The 2nd light-receiving means that receives the 2nd light volume that does not contain said catoptric light while said luminescent means is not emitting light,

A means to generate the 3rd picture that includes depth information in each pixel value by deducting said 2nd light volume from said 1st light volume, and extracting an ingredient of said catoptric light from said 1st light volume,

A means to compute depth information corresponding to each field of said 3rd picture to said 1st picture,

It provides,

The terminal unit according to claim 1, wherein said extraction means extracts three-dimensional physical relationship three-dimensional shape for [ said ] photography, and for [ said / two or more ] photography from said depth information corresponding to each field of said 1st picture and the 1st picture concerned at least.

[Claim 12]

The terminal unit according to claim 10 or 11 acquiring said 3rd picture simultaneously with said 1st picture.

[Claim 13]

The terminal unit according to claim 1, wherein a picture acquired by said 1st acquisition means is a color picture.

[Claim 14]

Said creating means generates said 2nd picture by compounding an added picture which expressed effect expressions and objects, such as a shock, in said 1st picture,

A memory measure which memorizes two or more kinds of added pictures for compounding in said 1st picture by this creating means,

A means to choose an added picture compounded in said 1st picture from two or more added pictures memorized by this memory measure,

The terminal unit according to claim 1 providing in a pan.

[Claim 15]

Said creating means generates said 2nd picture by compounding an added picture which expressed effect expressions and objects, such as a shock, in said 1st picture,

A memory measure which memorizes two or more kinds of added pictures for compounding in said 1st picture by this creating means,

A means to extract at least one of an object reflected to the 1st picture concerned, color atmosphere, and composition of a picture from said 1st picture as a feature of a scene,

A means to choose an added picture compounded in said 1st picture based on the feature of said scene from two or more kinds of added pictures memorized by said memory measure,

The terminal unit according to claim 1 providing in a pan.

[Claim 16]

The terminal unit according to claim 1 providing further a sound effect addition means which adds a sound effect suitable for said 2nd picture.

[Claim 17]

A displaying means which displays said 2nd picture,

A means of communication which transmits and receives said 2nd picture at least between terminal units of others of said two or more terminal units,

The terminal unit according to claim 1 providing in a pan.

[Claim 18]

It is said server apparatus in a communications system which comprises a server apparatus mutually connected with a terminal unit of plurality which can communicate, and a terminal unit of the plurality concerned so that communication was possible, A reception means which receives depth information corresponding to each of two or more fields which consist of 1 or two or more pixels which constitute the 1st picture that was transmitted from one of said two or more terminal units, and containing two or more candidates for photography, and the 1st picture concerned,

An extraction means to extract three-dimensional physical relationship three-dimensional shape for [ said / two or more ] photography, and for [ said / two or more ] photography at least based on said 1st picture received by this reception means, and said depth information,

A creating means which generates the 2nd picture by compounding an advertisement image in said 1st picture based on at least one of three-dimensional shape for [ which were acquired by this extraction means / said / two or more ] photography, and three-dimensional physical relationship,

A transmitting means which transmits said 2nd picture to a terminal unit of others of said two or more terminal units,

A providing server apparatus.

[Claim 19]

It is said server apparatus in a communications system which comprises a server apparatus mutually connected with a terminal unit of plurality which can communicate, and a terminal unit of the plurality concerned at least so that communication was possible,

The 1st picture that was transmitted from one of said two or more terminal units and that contains two or more candidates for photography at least, A reception means which was extracted from depth information corresponding to each of two or more fields which consist of 1 or two or more pixels which constitute the 1st picture concerned and which receives at least three-dimensional shape for [ said / two or more ] photography, and characteristic information including three-dimensional physical relationship for [ said / two or more ] photography,

A creating means which generates the 2nd picture by compounding an advertisement image in said 1st picture based on said characteristic information received by this reception means,

A transmitting means which transmits said 2nd picture to a terminal unit of others of said two or more terminal units,

A providing server apparatus.

[Claim 20]

While acquiring the 1st picture containing two or more candidates for photography, depth information corresponding to each of two or more of said fields of said 1st picture that comprises two or more fields which consist of 1 or two or more pixels is acquired,

An image work method extracting three-dimensional physical relationship three-dimensional shape for [ said / two or more ] photography, and for [ said / two or more ] photography as characteristic information at least, and processing said 1st picture based on this characteristic information based on said 1st picture and said depth information.

---

## DETAILED DESCRIPTION

---

[Detailed Description of the Invention]

[0001]

[Field of the Invention]

This invention about the image work method which combines a desired picture in the photoed picture, for example in terminal units, such as a cellular phone, and processes a picture. It is related with the terminal unit and communications system which communicate by transmitting and receiving the picture especially generated using this image work method between the terminal units of plurality, such as a cellular phone.

[0002]

[Description of the Prior Art]

In recent years, the camera devices (video camera etc.) which can photo video generally came to spread widely. This became possible to photo and peruse the familiar subject of events, such as one's child's growth process, and an athletic meet, in comfort. The camera possesses in a cellular phone, PDA, etc. in many cases, video can be photoed freely in a street corner etc., it can be with an E-mail etc. these days as it is, and the mail which attached video to the partner can be sent now. By these, the utilizing method which uses a video image as a means of communication is gaining popularity (for example, refer to the nonpatent literature 1 and nonpatent literature 2). Network broadband-ization also progresses quickly and image communication of real time of a TV phone etc. is also already realized (for example, refer to nonpatent literature 3).

[0003]

As mentioned above, the culture [ general public ] using video as communication in comfort is permeating. However, in the conventional video camera, it is most to use as a raw material what was simply reflected to the camera, and giving a certain special effects was seldom performed to it. This is because it was that in which the conventional camera photos only the target sexual desire news in two dimensions unlike human being's eyes. Therefore, in human being's eyes, what was reflected to the camera can attach distinction only by sexual desire news to the ability to grasp distinction of the object reflected to the camera, the context, cubic shape, etc.

[0004]

For this reason, it was dramatically difficult to perform processing in which take out only human being's portion, and distinguish from a background or the shape change of the surface of OK and a thing is distinguished for two or more human beings' context, and physical relationship for example, from the picture photoed with the camera. The trial in which such information will be acquired virtually also occurs only using sexual desire news. For example, the object out of a picture is separated from a background, and the art of performing picture communication only using an object is indicated using this (for example, refer to patent documents 1). Here, parts, such as a face, were separated and an effect which is communicating in cyber-space is given by displaying two or more persons' face parts which are having a dialog on the background prepared independently. However, since sexual desire news unrelated to three-dimensional information is originally used, it is difficult for restrictions of many [ grasp / the context of an object that it is difficult to distinguish an object strictly and actual ] to exist, and to carry out stably.

[0005]

So, in the communication using the present video image, applying special effects, such as picture composition, to the photoed video image using information, including physical relationship, such

as distinction of such an object and a context, and cubic shape, was not considered. The special effects that being carried out now carries out overwrite of the character on a video image, or it piles up a decoration frame are almost the case (for example, refer to nonpatent literature 4).

[0006]

On the other hand, the image communication between the above-mentioned individuals is extended, and if it becomes possible to show an advertisement into the image, various business deployment can be considered. However, as mentioned above, when advertising addition is considered in the conventional picture, there is a problem how an advertisement should be superimposed. If the advertisement was only superimposed as it was on the conventional picture, it becomes obstructive in many cases. For example, when the superimposed display of the advertisement is carried out on the face of itself or a partner in the midst of carrying out communication by a TV phone using human being's bustup picture, it not only becomes obstructive, but a user has a bad impression to the advertisement, and there is no advertising meaning. Then, the trial in which an advertisement was conventionally added into a picture was not performed.

[0007]

[Patent documents 1]

JP,2001-188910,A gazette (a paragraph number "0158" thru/or "0170", Drawing 24)

[0008]

[Nonpatent literature 1]

The WWW page of J-Phone

<http://www.j-phone.com/movie-shamail/>

[0009]

[Nonpatent literature 2]

The WWW page of NTT DoCoMo

[http://www.nttdocomo.co.jp/p\\_s/imode/ishot/index.html](http://www.nttdocomo.co.jp/p_s/imode/ishot/index.html)

[0010]

[Nonpatent literature 3]

The WWW page of NTT DoCoMo FOMA

<http://foma.nttdocomo.co.jp/>

[0011]

[Nonpatent literature 4]

The WWW page of ATLUS

<http://www.atlus.co.jp/am/printclub/>

[0012]

[Problem(s) to be Solved by the Invention]

Thus, when combining the picture of other requests in the photoed picture (video and a still picture are included) and processing a picture conventionally, two or more three-dimensional shape, physical relationship for photography, etc. of each in the photoed picture concerned are disregarded, The problem that it could only perform processing combining a picture simply etc. was on the two-dimensional flat surface on a picture.

[0013]

However, if two or more candidates for photography in the photoed picture concerned can be distinguished and information, including those three-dimensional shape, physical relationship, etc., can be used, The character of computer graphics (CG) is flying about around human being who utilized the three-dimensional information for [ in these pictures ] photography, and has

been reflected into the photoed video concerned, or, As it said changing only background parts, it becomes possible to give the special effects which were more varied.

[0014]

If distinction of the portions of a background and a person is attained, when compounding an advertisement image, it also becomes possible to compound an advertisement image to background parts so that an advertisement may not be applied to human being's portion. It becomes realizable [ the presenting method of the new advertisement which employed the target cubic shape efficiently well ] that the character of the computer graphics which had an advertisement in the surroundings of human being reflected into the photoed video flies about etc.

[0015]

Then, in view of the above-mentioned problem, this invention by conventional technology, [ impossible ] It doubles with the three-dimensional features, such as each three-dimensional shape for [ in the photoed picture (video and a still picture are included) / two or more ] photography, and three-dimensional physical relationship, It aims at providing the image processing unit, communication terminal device, and server apparatus using the image work method and it which can process combining an advertisement image and a desired picture etc.

[0016]

[Means for Solving the Problem]

(1) The 1st acquisition means that acquires the 1st picture that this invention is a terminal unit which is one of two or more terminal units which can communicate mutually, and contains two or more candidates for photography, The 2nd acquisition means that said 1st picture comprises two or more fields which consist of 1 or two or more pixels, and acquires depth information corresponding to each of two or more of these fields, An extraction means to extract three-dimensional physical relationship three-dimensional shape for [ said / two or more ] photography, and for [ said / two or more ] photography as characteristic information at least based on said 1st picture and said depth information, By having provided a creating means which processes said 1st picture and generates the 2nd picture based on characteristic information extracted by this extraction means, According to the three-dimensional features, such as each three-dimensional shape for [ in the 1st photoed picture (video and a still picture are included) / two or more ] photography, and three-dimensional physical relationship, it is processible to combine a desired picture etc.

[0017]

For example, said creating means is what generates said 2nd picture by combining an objective picture in said 1st picture, In that case, a picture of the object concerned is combined in the 1st picture concerned based on at least one of three-dimensional physical relationship of for concerned an object and for [ said / two or more ] photography when a position of the object concerned is provided in a depth direction in the 1st picture concerned, and three-dimensional shape for [ said / two or more ] photography.

[0018]

By combining an objective picture in said 1st picture, said creating means generates said 2nd picture, in that case, it controls it so that a motion of said object suits three-dimensional shape for [ said ] photography, and it combines a picture of the object concerned in the 1st picture concerned.

[0019]

Extract said extraction means and a position of a depth direction where said candidate for



photography in said 1st picture exists, and a background region of the back for [ concerned ] photography said creating means, By combining an objective picture in said 1st picture, said 2nd picture is generated, and in that case, a picture of the object concerned is combined in the 1st picture concerned so that the object concerned may exist between said background regions for photography concerned.

[0020]

The 1st position of a depth direction where the 1st candidate for photography in which said extraction means is one of for [ in said 1st picture / said / two or more ] photography exists, Extract the 2nd position of a depth direction where the 2nd candidate for photography that is one of everything [ the ] but inside for [ in the back 1st for / concerned / photography / said / two or more ] photography exists, and said creating means, By combining an objective picture in said 1st picture, said 2nd picture is generated, and in that case, a picture of the object concerned is combined in the 1st picture concerned so that the object concerned may exist between the 1st position concerned and the 2nd position.

[0021]

Extract said extraction means as a three-dimensional position in which said candidate for photography in said 1st picture exists, and the 3rd position said creating means, When said 2nd picture is generated and the object concerned results in said 3rd position based on a motion of the object concerned and a motion for [ concerned ] photography at that time by combining an objective picture in said 1st picture, It judges that the object concerned collided with the candidate for photography concerned, it controls so that a motion and expression of the object concerned correspond to a collision, a picture of the object concerned is combined in the 1st picture concerned, and when it judges that said object collided with said candidate for photography, effect expression of a collision is compounded in said 1st picture.

[0022]

Extract said extraction means as a three-dimensional position in which said candidate for photography in said 1st picture exists, and the 3rd position said creating means, By combining an objective picture in said 1st picture, generate said 2nd picture and in that case Said 3rd position for [ concerned ] photography, Based on the 4th position of being a three-dimensional position of the object concerned in said 1st picture, a motion and expression of the object concerned are controlled and a picture of the object concerned is combined in the 1st picture concerned.

[0023]

Extract said extraction means from said 1st picture, and each picture for [ said / two or more ] photography, and a picture of a background region after them said creating means, Said 2nd picture is generated by processing at least one of a picture of said background region, and each pictures for [ said / two or more ] photography.

[0024]

(2) A means to acquire the 3rd picture that said 2nd acquisition means corresponds to said 1st picture, and includes depth information in each pixel value, Provide a means to extract depth information corresponding to each field of said 3rd picture to said 1st picture, and said extraction means, Three-dimensional physical relationship three-dimensional shape for [ said ] photography and for [ said / two or more ] photography is extracted from said depth information corresponding to each field of said 1st picture and the 1st picture concerned at least.

[0025]

(3) A luminescent means in which said 2nd acquisition means emits light towards said candidate for photography, While the 1st light-receiving means that receives the 1st light volume

containing catoptric light from the candidate for photography concerned of light made this luminescent means emit light and irradiated said candidate for photography, and said luminescent means are not emitting light, The 2nd light-receiving means that receives the 2nd light volume that does not contain said catoptric light, and by deducting said 2nd light volume from said 1st light volume, and extracting an ingredient of said catoptric light from said 1st light volume, Provide a means to generate the 3rd picture that includes depth information in each pixel value, and a means to compute depth information corresponding to each field of said 3rd picture to said 1st picture, and said extraction means, Three-dimensional physical relationship three-dimensional shape for [ said ] photography and for [ said / two or more ] photography is extracted from said depth information corresponding to each field of said 1st picture and the 1st picture concerned at least.

[0026]

(4) Said creating means is what generates said 2nd picture by compounding an added picture which expressed effect expressions and objects, such as a shock, in said 1st picture, A memory measure which memorizes two or more kinds of added pictures for compounding in said 1st picture by this creating means, By having provided further a means to choose an added picture compounded in said 1st picture from two or more added pictures memorized by this memory measure, the user can choose a desired added picture freely and can add special effects which suited liking of them to the 1st picture of the above.

[0027]

(5) Said creating means is what generates said 2nd picture by compounding an added picture which expressed effect expressions and objects, such as a shock, in said 1st picture, A memory measure which memorizes two or more kinds of added pictures for compounding in said 1st picture by this creating means, A means to extract at least one of an object reflected to the 1st picture concerned, color atmosphere, and composition of a picture from said 1st picture as a feature of a scene, Special effects suitable for an object in a picture of the above 1st, color atmosphere, composition of a picture, etc. can be added by having provided further a means to choose an added picture compounded in said 1st picture from two or more kinds of added pictures memorized by said memory measure, based on the feature of said scene.

[0028]

(6) This invention is said server apparatus in a communications system which comprises a server apparatus mutually connected with a terminal unit of plurality which can communicate, and a terminal unit of the plurality concerned so that communication was possible, The 1st picture that was transmitted from one of said two or more terminal units and containing two or more candidates for photography, A reception means which receives depth information corresponding to each of two or more fields which consist of 1 or two or more pixels which constitute the 1st picture concerned, An extraction means to extract three-dimensional physical relationship three-dimensional shape for [ said / two or more ] photography, and for [ said / two or more ] photography at least based on said 1st picture received by this reception means, and said depth information, A creating means which generates the 2nd picture by compounding an advertisement image in said 1st picture based on at least one of three-dimensional shape for [ which were acquired by this extraction means / said / two or more ] photography, and three-dimensional physical relationship, a picture (the 1st picture of the above) transmitted and received between terminal units by having provided a transmitting means which transmits said 2nd picture to a terminal unit of others of said two or more terminal units -- the 1st picture (video.) concerned While a still picture is included, according to the three-dimensional features,

such as each three-dimensional shape for [ two or more ] photography, and three-dimensional physical relationship, an advertisement image (it does not become the obstacle for [ in the 1st picture concerned ] photography like) is compoundable. Therefore, an advertisement image can be compounded efficiently and improvement in advertising effectiveness can be aimed at.

[0029]

The 1st picture to which the above-mentioned server apparatus was transmitted from one of said two or more terminal units and that contains two or more candidates for photography at least, . Were extracted from depth information corresponding to each of two or more fields which consist of 1 or two or more pixels which constitute the 1st picture concerned. A reception means which receives at least characteristic information including three-dimensional physical relationship three-dimensional shape for [ said / two or more ] photography, and for [ said / two or more ] photography, Based on said characteristic information received by this reception means, it may comprise a creating means which generates the 2nd picture, and a transmitting means which transmits said 2nd picture to a terminal unit of others of said two or more terminal units by compounding an advertisement image in said 1st picture.

[0030]

[Embodiment of the Invention]

Hereafter, an embodiment of the invention is described with reference to drawings.

[0031]

(A 1st embodiment)

First, a 1st embodiment of this invention is described.

[0032]

<The whole composition>

Drawing 1 shows the example of composition of the principal part concerning a 1st embodiment which is one of two or more terminal units which can communicate mutually.

[0033]

The picture communication device shown in drawing 1, The image acquiring part 1 which acquires the depth information of the picture for photography, and the picture concerned, The feature extraction part 2 which extracts the three-dimensional features for [ in the picture concerned ] photography, etc. based on the depth information of the picture acquired by the image acquiring part 1, Based on the feature extracted by the feature extraction part 2, a three-dimensional element is taken into consideration in the picture acquired by the image acquiring part 1, It comprises the picture presentation part 4 which presents the processing section 3 which gives special effects, such as picture composition, the picture acquired as a result of processing it by the processing section 3, and the picture received from the communications department 5, and the communications department 5 for transmitting and receiving the picture processed by the processing section 3.

[0034]

Among two or more terminal units which have composition as shown in drawing 1, it communicates among the users of each terminal unit by transmitting the picture acquired by processing it by the processing section 3 to the picture acquired in each to other terminal units, or receiving.

[0035]

<Image acquiring part>

First, the image acquiring part 1 is explained.

[0036]

While the image acquiring part 1 acquires a color picture as a picture for photography, The depth information of the picture concerned is acquired and it acquires as a color picture (here, it is called a depth color picture) including the depth information which reflected the three-dimensional shape and distance from the image acquiring part 1 to the candidate for photography concerned for the candidate for photography.

[0037]

Although explained here taking the case of the case where a color picture is acquired as a picture for photography, it may not restrict in this case but they may be gradation images (what is called a monochrome image). In this case, in the image acquiring part 1, the candidate for photography is acquired as a depth monochrome image reflecting that three-dimensional shape and distance from the image acquiring part 1 to the candidate for photography concerned.

[0038]

Now, as the color picture comprises a unit called a pixel and it is shown in drawing 6 (b) at a pixel, the information on the color of RGB is usually stored. It realizes by arranging this pixel in array form in the two-dimensional direction in every direction. For example, the color picture of VGA (Video Graphics Array) size is expressed by the two dimensional array of 640 pixels of the directions of a x axis (width), and 480 pixels of the directions of the y-axis (length), and the information on the color in the position is stored for example, in RGB form in each pixel.

Drawing 6 (b) shows an example of the information on the color stored in each pixel, including for example, = (R, G, B) (r1, g1, b1) etc., in the color picture of 5 pixels of x shaft orientations, and 8 pixels of y shaft orientations.

[0039]

In addition to the information on the color of RGB form, in the depth color picture, depth information (for example, distance information from image acquiring part 1 to candidate for photography) d in the position is matched with each pixel as mentioned above, for example. As matching here, to each pixel of a depth color picture, depth information d corresponding to the pixel concerned shall be further added to the above (R, G, B), and it shall memorize in the form (R, G, B, d), for example.

[0040]

In using a monochrome image instead of a color picture, the pixel value of the monochrome image concerned serves as a gradient. Therefore, depth information d corresponding to the pixel concerned is further memorized by each pixel of the depth monochrome image in this case at the above-mentioned gradient.

[0041]

Although explained that each pixel of a depth color picture holds depth information d, it does not restrict in this case, but the field which consists of two or more pixels beforehand is appointed as a unit field, and it may be made to match or hold depth information d for every unit field of this here.

[0042]

If this depth information is not used when displaying a depth color picture though depth information was added to sexual desire news, such as RGB, at each pixel of the depth color picture, a depth color picture is displayed completely like the usual color picture. A color picture can also be displayed in three dimensions by using the depth information corresponding to each pixel.

[0043]

Comparison of the usual monochrome image and a depth monochrome image is shown in

drawing 2. Drawing 2 (a) is the usual monochrome image. It is the picture to which the gradient is dedicated by each pixel in two dimensions. An example of a depth monochrome image is shown in drawing 2 (b). This is expressed in three dimensions using depth information. (It was made to see from the viewpoint from under a transverse plane so that intelligibly [ a three-dimensional expression ]) In this way a depth monochrome image and a depth color picture, It is the feature that the distance information  $d$  from the image acquiring part 1 to the candidate for photography is included as depth information unlike a conventional monochrome image and color picture.

[0044]

What matched depth information with each pixel of the usual color picture and the monochrome image here, Although it is also called a color picture including depth information and a monochrome image including depth information, respectively, a depth color picture and a depth monochrome image are in positioning called one example of a color picture including depth information and a monochrome image including depth information, respectively.

[0045]

Then, the composition of the image acquiring part 1 is explained in detail using drawing 3. The picture information acquisition part 101 for acquiring the picture information in which drawing 3 is a functional block diagram of the image acquiring part 1, is roughly divided into, and the sexual desire news for photography is contained in real time, It comprises the depth information acquisition part 102 for acquiring the depth information for photography in real time, the imaging operation control section 104, the depth color picture generation part 113, and the outputting part 114.

[0046]

The natural optical image image pick-up part 103 which acquires the natural optical image which is a color picture of the object which includes a background when the picture information acquisition part 101 picturizes the candidates for photography (for example, object etc.) by available light (the illumination light is included), The natural optical image storage parts store 105 in which the natural optical image picturized in the natural optical image image pick-up part 103 is stored, The natural optical image controller 106 which reads the natural optical image stored in the natural optical image storage parts store 105, and adjusts brightness, contrast, etc. if needed, The natural optical image which comprised the picture information outputting part 107 outputted to the processing section 3 which outputs the natural optical image adjusted by the natural optical image controller 106 in a data format suitable for processing by the processing section 3, and was adjusted by the natural optical image controller 106 is further outputted to the depth color picture generation part 113.

[0047]

The natural optical image image pick-up parts 103 are CCD and a CMOS image sensor, and acquire the color plane pictures (for example, object etc.) for [ including a background ] photography in real time, for example. Thereby, a color is acquirable among objective attributes. Since the above-mentioned color picture is picturized continuously, objective motion information can be grasped from change of a continuous frame. An objective picture can be used as an animation.

[0048]

The imaging operation control section 104 generates the control signal for controlling operation of the natural optical image image pick-up part 103 and the catoptric light picture image pick-up part 108.

[0049]

The catoptric light picture image pick-up part 108 for the depth information acquisition part 102 to, picturize the catoptric light picture for [, such as a photographic subject in an object or its circumference, ] photography on the other hand according to the control signal from the imaging operation control section 104, The catoptric light image storage section 109 in which the catoptric light picture picturized in the catoptric light picture image pick-up part 108 is stored, The catoptric light picture amendment part 110 which reads the catoptric light picture stored in the catoptric light image storage section 109, and carries out various amendments, The parameter storing part 111 the parameter for amending a catoptric light picture is remembered to be, The depth information which the depth information operation part 112 which analyzes the catoptric light picture amended in the catoptric light picture amendment part 110, and calculates objective depth information was consisted of, and was calculated by the depth information operation part 112 is outputted to the depth color picture generation part 113.

[0050]

Here, the details of the catoptric light picture image pick-up part 108 are explained with reference to drawing 4. Drawing 4 is a functional block diagram of the catoptric light picture image pick-up part 108. The catoptric light picture image pick-up part 108 comprises the light-emitting part 114, the light-receiving optical system 115, the catoptric light extraction part 116, and the timing control section 117.

[0051]

The light-emitting part 114 generates the light which carries out intensity fluctuation in time according to the timing signal generated by the timing control section 117. This light is ahead of the light-emitting part 114, for example, it is irradiated with it by a user's head (an objective example) by drawing 4. Although it is a near infrared as a light generated here, not only this but lights of other wavelength areas, such as visible light, can be used.

[0052]

For example, it is condensed by the light-receiving optical system 115 which comprises a lens etc., and image formation of the catoptric light from a user's head is carried out on the acceptance surface of the catoptric light extraction part 116. The filter which passes a near infrared is formed in the light-receiving optical system 115. Outdoor daylight like visible light other than a near infrared or far-infrared light is cut among catoptric light with this filter.

[0053]

The catoptric light extraction part 116 extracts the spatial intensity distribution of the catoptric light which forms the above-mentioned image formation. This intensity distribution can be regarded as a picture by catoptric light. This is the above-mentioned catoptric light picture (silhouette image). In order to attain this function, the catoptric light extraction part 116 comprises the 1st light sensing portion 119, 2nd light sensing portion 120, and difference operation part 121. The 1st light sensing portion 119 and 2nd light sensing portion 120 receive light to different timing. And the timing control section 117 controls such operation timing so that the light-emitting part 114 does not emit light, while the light-emitting part 114 emits light while the 1st light sensing portion 119 is receiving light, and the 2nd light sensing portion 120 is receiving light. Thereby, the 1st light sensing portion 119 receives the catoptric light by the object of the light from the light-emitting part 114, and the other available light (that is, outdoor daylight, such as sunlight and illumination light). On the other hand, the 2nd light sensing portion 120 receives only available light. Although the timing which both receive differs, since it is near, change of the available light in during this period can be disregarded.

[0054]

Therefore, if the difference of the image which received light by the 1st light sensing portion 119, and the image which received light by the 2nd light sensing portion 120 is taken, only the ingredient of the catoptric light by an object is extracted among the lights from the light-emitting part 114, and a catoptric light picture can be generated. The difference operation part 121 calculates and outputs the difference of the image which received light by the 1st light sensing portion 119 and 2nd light sensing portion 120.

[0055]

The catoptric light from an object decreases substantially as the distance of an object and the catoptric light picture image pick-up part 108 becomes long. When the objective surfaces are uniformly scattered about in light, the light income per 1 pixel of catoptric light pictures becomes small in inverse proportion to the square of the distance to an object. That is, when the pixel value of the pixel in the coordinates (i, j) in a catoptric light picture is set to Q (i, j) for example,  $Q(i, j) = K/d^2$  -- (1)

It can express. Here, K is the coefficient adjusted at the time of  $d = 0.5$  m, for example so that the value of Q (i, j) might be set to "255." Distance can be found by solving a formula (1) about d.

[0056]

Thus, the intensity value of catoptric light convertible into the distance from the catoptric light picture image pick-up part 108 to the candidate for photography, i.e., depth information, is included in each pixel of the catoptric light picture as depth information.

[0057]

Since the intensity value of the catoptric light from an object is convertible into the distance from the catoptric light picture image pick-up part 108 to an object, it can grasp objective cubic shape. The catoptric light from a background is small to the extent that it can be disregarded mostly. Therefore, the catoptric light picture of the photographic subject in an object and its circumference into which the background was cut can be acquired.

[0058]

An example of a catoptric light picture is shown in drawing 6 (a). Since it is easy, the case of the 5x8-pixel catoptric light picture which is a part of 256x256-pixel catoptric light picture is shown in drawing 6 (a). The pixel value of each pixel is an intensity value of catoptric light.

[0059]

The three-dimensional image of the hand in the catoptric light picture concerned produced from the above-mentioned depth information included in the catoptric light picture photoed in the catoptric light picture image pick-up part 108 by making human being's hand applicable to photography is shown in drawing 5.

[0060]

A catoptric light picture can be acquired with the device which has the following composition of (1) to (3), for example. (1) The timing signal creating means for generating the pulse signal and modulating signal which change uniformly or in time in time. (2) The luminescent means for emitting the light which carries out intensity change based on the signal generated by this timing signal creating means. (3) A catoptric light extraction means to arrange and constitute a means to separate and detect the catoptric light by the object of the light emitted from this luminescent means from outdoor daylight (available light) synchronizing with the signal from a timing signal creating means, and to detect the catoptric light picture by the object of light.

[0061]

Still more detailed explanation of the above-mentioned component of the catoptric light picture

image pick-up part 108 is indicated to JP,10-177449,A for which the same applicant as this case applied.

[0062]

Since the catoptric light picture image pick-up part 108 picturizes continuously the catoptric light picture of the photographic subject containing an object, it can grasp objective motion information from change of a continuous frame. Thereby, an objective picture can be used as an animation.

[0063]

A depth color picture is generated combining the depth information acquired by calculating a natural optical image (plane picture of the color of an object including a background), and an objective catoptric light picture so that it may mention later. For this reason, as for the catoptric light picture image pick-up part 108 and the natural optical image image pick-up part 103, since it is necessary to make the pixel of a catoptric light picture, and the pixel of a natural optical image correspond, approaching and arranging is preferred. It is still more preferred these natural optical image image pick-up parts 103 from this viewpoint, the catoptric light picture image pick-up part 108, and that 1 chip making of the imaging operation control section 104 is carried out to necessity if needed.

[0064]

Between a natural optical image and a catoptric light picture, it is required to make matching between the fields which consist between pixels or of two or more pixels, for example. For example, suppose that the natural optical image (color picture) acquired by the image acquiring part 101 is a picture as shown in drawing 6 (b), and it is a picture as the catoptric light picture acquired by the depth information acquisition part 102 showed to drawing 6 (a). In this case, it is matched with 1 to 1 between each pixel of a color picture, and each pixel of a catoptric light picture. When each pixel in each picture of drawing 6 (a) and drawing 6 (b) is expressed using the coordinates (i, j) (here, it is  $i = 1-5$  and  $j = 1-8$ ), the pixel P2 (i, j) of the catoptric light picture in drawing 6 (a) and the pixel P1 (i, j) in the natural optical image of drawing 6 (b) turn into a mutually corresponding pixel. For example, it becomes the pixel P2 (5, 8) of the catoptric light picture in drawing 6 (a), and a pixel corresponding mutually [ the pixel P1 (5, 8) in the natural optical image of drawing 6 (b) ]. Therefore, the distance (depth information) d by which the pixel value of the pixel P3 (i, j) in a depth color picture was computed from the sexual desire news of the pixel P1 (i, j) in a natural optical image and the pixel P2 (i, j) of the catoptric light picture (with depth information operation part 112) is included.

[0065]

Although the case where the pixel of a natural optical image and the pixel of the catoptric light picture were matched was shown here, In this case, do not restrict but from the difference in the size of a natural optical image and a catoptric light picture, resolution, etc. For example, a natural optical image is divided into two or more unit fields which consist of two or more pixels, It may be made to match between each of two or more of these unit fields, and each pixel of a catoptric light picture, and a catoptric light picture is conversely divided into two or more unit fields which consist of two or more pixels, and it may be made to match between each of two or more of these unit fields, and each pixel of a natural optical image. Each of a catoptric light picture and a natural optical image is divided into two or more unit fields which consist of two or more pixels, and it may be made to match between each unit field of a catoptric light picture and a natural optical image.

[0066]



When the image pick-up directions from a user are inputted into the imaging operation control section 104, for example, A control signal is outputted to the natural optical image image pick-up part 103 and the catoptric light picture image pick-up part 108 so that a natural optical image and a catoptric light picture may be acquired almost simultaneous in the natural optical image image pick-up part 103 and the catoptric light picture image pick-up part 108.

[0067]

Next, amendment in the catoptric light picture amendment part 110 shown in drawing 3 is explained. It amends in consideration of an objective color, a reflection property, etc. If it explains in detail, the intensity of the catoptric light from an object will be influenced by factors other than the distance of an object and the catoptric light picture image pick-up part 108. For this reason, even if it finds distance simply from a catoptric light picture, distance (that is, cubic shape) is not sometimes exact. For example, when the color of an object surface is black, the intensity of catoptric light falls. When the objective surface contains many mirror reflection components, strong catoptric light occurs in the portion to which the normal of the objective surface becomes close to the direction of a light source.

[0068]

Therefore, in the catoptric light picture amendment part 110 before searching for distance information from objective reflected light intensity, The natural optical image which refers to the parameter about an objective surface color, a reflection property, etc. which are beforehand stored in the parameter storing part 111, or is stored in the natural optical image storage parts store 105 is referred to, An objective catoptric light picture is amended based on the pixel value of the field (a pixel and a unit field) corresponding to the pixel and unit field of a catoptric light picture in the natural optical image concerned. For example, if the sexual desire news of the pixel of the natural optical image corresponding to a certain pixel in a catoptric light picture is "black", the pixel value concerned of a catoptric light picture will be amended using the parameter at the time of receiving the catoptric light from the black candidate for photography.

[0069]

Since each pixel value of the catoptric light picture acquired in the catoptric light picture image pick-up part 108 is amended by this, the accuracy of the depth information computed by the latter depth information operation part 112 can be raised.

[0070]

About each pixel of the catoptric light picture amended in the catoptric light picture amendment part 110, the depth information operation part 112 will find the distance  $d$  corresponding to the pixel concerned, for example from a following formula (2), if light income (reflected light intensity) as the pixel value is set to  $Q$ .

[0071]

$$d=(K/Q)^{1/2} \text{ -- (2)}$$

That is, since the intensity value of the catoptric light from an object is convertible into the distance  $d$  from the catoptric light picture image pick-up part 108 to an object as mentioned above, this distance  $d$  is found as depth information.

[0072]

Although the case where the distance  $d$  as depth information was found was shown about each pixel in a catoptric light picture, it does not restrict in this case but may be made to compute the distance  $d$  for every above-mentioned unit field here from the average value of the pixel value in that representative picture element or the unit field concerned, etc., for example.

[0073]

Now, in the depth color picture generation part 113, a depth color picture is generated from the depth information of each pixel (or each unit field) in the natural optical image outputted from the natural optical image controller 106, and the catoptric light picture computed by the depth information operation part 112.

[0074]

That is, the depth information computed from the pixel of the catoptric light picture corresponding to the pixel concerned by the pixel value of each pixel of a color picture as shown in drawing 6 (b) which is a natural optical image, for example is added, and each pixel value of a depth color picture is generated. For example, the pixel P2 (5, 8) of the catoptric light picture in drawing 6 (a) and the pixel P1 (5, 8) in the natural optical image of drawing 6 (b) turn into a mutually corresponding pixel. Therefore, the distance (depth information)  $d$  computed from the pixel P2 (5, 8) of the catoptric light picture (with depth information operation part 112) is added to the pixel value of the pixel P1 (5, 8) in a natural optical image, and the pixel value of the pixel P3 (5, 8) of a depth color picture is generated to it.

[0075]

When dividing a natural optical image into two or more unit fields which consist of two or more pixels and matching between each of two or more of these unit fields, and each pixel of a catoptric light picture, Depth information  $d$  computed from the pixel of the catoptric light picture corresponding to the unit field concerned by each unit field of the natural optical image is matched. Each pixel value of the depth color picture (to which for example, depth information  $d$  computed by the pixel value of each pixel in the unit field concerned from the pixel of the catoptric light picture corresponding to the unit field concerned is added) is generated.

[0076]

When dividing a catoptric light picture into two or more unit fields which consist of two or more pixels conversely and matching between each of two or more of these unit fields, and each pixel of a natural optical image, The depth information computed from the unit field of the catoptric light picture corresponding to the pixel concerned by the pixel value of each pixel of a natural optical image is added, and each pixel value of a depth color picture is generated.

[0077]

When dividing each of a catoptric light picture and a natural optical image into two or more unit fields which consist of two or more pixels and matching between each unit field of a catoptric light picture and a natural optical image, Depth information  $d$  computed by each unit field of the natural optical image from the unit field of the catoptric light picture corresponding to the unit field concerned is matched. Each pixel value of the depth color picture (to which for example, depth information  $d$  computed by the pixel value of each pixel in the unit field concerned from the unit field of the catoptric light picture corresponding to the unit field concerned is added) is generated.

[0078]

Thus, the generated depth color picture is sent to the outputting part 114, in order to double with the protocol and data format of an output destination change, it performs conversion etc. here, and it is outputted to the feature extraction part 2.

[0079]

The composition of the image acquiring part 1 explained above is an example to the last, and is not limited to this. When acquiring depth information especially, it is not necessary to necessarily use a catoptric light picture as mentioned above. . Namely, say that depth information is calculated by using the parallax information of the picture photoed from two or more views. It

may be the composition of acquiring depth information using the technique of stereo matching, and it irradiates with a stripes-like laser beam and the method called the laser range finder of measuring the above-mentioned depth information using distortion of the form may be used. What can acquire depth information using methods other than these, and can acquire the above depth color pictures can also be used.

[0080]

Although the described image acquisition part 1 generates the picture (depth color picture) which includes depth information in each pixel by the depth color picture generation part 113, In this case, it may not restrict, but the depth information corresponding to the pixel concerned may be matched with each pixel of a natural optical image, and that correspondence relation may only be held or memorized.

[0081]

<Feature extraction part>

Next, the feature extraction part 2 is explained. Here, let the depth color picture which includes depth information in the pixel value calculated by the image acquiring part 1 be a processing object.

[0082]

The feature extraction part 2 is for extracting the three-dimensional feature for photography based on the depth information included in the depth color picture acquired by the image acquiring part 1.

[0083]

Here, the three-dimensional feature analyzed by the feature extraction part 2 with reference to the depth color picture of the contents shown in drawing 7 is explained concretely. Since drawing 7 shows the case where it displays without using the depth information included in a depth color picture, it is the same as that of the usual color picture (in the case of drawing 7 monochrome image). The upper half of the body of the person by whom drawing 7 sat on the chair in the center of a scene is reflected, and the person mentioned the right arm and has upraised the index finger. Can juice is placed before the person.

[0084]

Drawing 8 showed typically the physical relationship of each object in this scene briefly. Like drawing 8, a can exists in the nearest distance from the image acquiring part 1 (refer to drawing 12), and the person exists beyond [ the ] (refer to drawing 10). Furthermore, there are background parts in the distance (refer to drawing 9). When the depth information within a person is looked at finely, a right arm portion is one of near portions from the image acquiring part 1 rather than the body (refer to drawing 11). That is, the difference in depth can divide a scene into some fields, as shown in drawing 9 - drawing 12.

[0085]

By thus, the thing for which the depth information included in the depth color picture acquired by the image acquiring part 1 is analyzed. What kind of candidate for photography (for example, object) exists in a scene, or (into what kind of field can it divide?). What has happened to the concavo-convex relation of the depth direction in a scene, or the features, such as three-dimensional physical relationship of each object (the portion which constitutes one object is also regarded as one object, respectively) in a scene, and cubic shape of each object, can be acquired.

[0086]

Then, the feature-analysis technique is explained.

[0087]

There is a thing of obtaining only the three-dimensional shape without the object currently picturized inside the scene distinguishing something as easiest analysis. For example, in the case of the scene of drawing 7, it is the middle-of-the-screen lower part (actually). The portion of a can has the nearest depth value, the screen left-hand side lower part (actually a person's right arm portion) has a depth value near the next, and it is the great portion of middle of the screen (actually). As it said that a person portion existed in the next and a subsequent portion (actually background parts) existed in the distance dramatically etc., the overall concavo-convex relation of a scene is extracted about a depth direction.

[0088]

In the feature extraction part 2, if necessary, it is also possible to conduct still more complicated analysis.

[0089]

Drawing 13 shows typically the depth information included in the depth color picture shown in drawing 7. Drawing 13 shows only the portion from which depth information is acquired as a solid line, seeing the scene shown in drawing 7 from the upper part. Thus, in a depth color picture, only the depth information of the portion which is visible from bearing of the exposure axis is acquired. (There is no data of the portion which is not visible from bearing of the exposure axis, for example, the portion hidden in this example behind the backside of a can or a person and a person's right arm etc.) The three-dimensional physical relationship for photography is acquired based on such depth information.

[0090]

The easiest method is distinguishing a foreground and a background simply by depth information d from the image acquiring part 1. A certain threshold TH is defined, and depth information considers that a portion nearer than the threshold TH is a foreground, and considers that a far portion is a background (refer to drawing 14). It is possible to distinguish what is reflected in the scene, and background parts by doing in this way. It can be distinguished whether any objects (that is, portion which hits a foreground) exist in the scene currently photoed by this.

[0091]

The scene which prepared two or more these thresholds, for example, has been reflected to the depth color picture with the 1st threshold x1, 2nd threshold x2, and the 3rd threshold x3. It may be made to identify whether from the distance x1, it divides into four fields like the distance x3 or subsequent ones between x3 with this side, and between x2 and distance x2, and an object exists about each field, and what the object is. [ the distance x1, and ]

[0092]

By what the break point (portion shown by the dotted line in drawing 15) of the depth direction called jump edge is detected for. As are shown in drawing 15, and it was called "background parts", the "person portion", "the person's right arm portion", the "portion of the can", etc., it is also possible to recognize the still finer physical relationship in a scene.

[0093]

Although various techniques can be used for the detecting method of jump edge, For example, the picture which extracted only the depth portion from the depth color picture (only depth information, such as a catoptric light picture, by the data located in a line with two-dimensional array form.) Henceforth, it can obtain by receiving calling it a depth picture and performing filtering processings (convolution filtering processing according to a Sobel operator typically etc.) for edge detection.

[0094]

Although there is the technique of registering the objective feature beforehand and looking for a similar portion with the feature in a picture of pattern matching, it can be recognized by using the technique what each object obtained above is. For example, if the candidates for photography "on the right of" etc. a "can", a "person", and a said person can be recognized and those physical relationship can moreover be recognized in the scene shown in drawing 7, from drawing 7. The complicated recognition like, such as there being "can", the position of the "can" being where, or a "person" being after a "can", and mentioning the "right hand", is also possible.

[0095]

The analytic method for extracting the three-dimensional feature from the above depth color pictures is an example to the last, and is not limited to this. It is possible to realize combining the technique of of other various analyses, image processing, and image recognition.

[0096]

In the feature extraction part 2, the three-dimensional feature in a natural optical image is extracted from a depth color picture (depth information acquired from each \*\* value of a natural optical image and a catoptric light picture). The three-dimensional feature in a natural optical image is the physical relationship (with the physical relationship of the plane direction in a natural optical image.) the three-dimensional shape for [ each ] photography (the uneven state on the surface is also included), and for [ two or more ] photography in a natural optical image, for example. The foreground part in which it is the physical relationship (mainly context) of the depth direction of a natural optical image, etc., and the candidate for photography distinguished from the position of the depth direction corresponding to each candidate for photography (based on the threshold defined beforehand) exists from these further, While being able to perform background parts and still finer area division, it can also be recognized with pattern matching etc. what the candidate for photography is.

[0097]

<Processing section>

Next, the processing section 3 is explained.

[0098]

The processing section 3 to the color picture (or depth color picture) which is a natural optical image acquired by the image acquiring part 1 based on the three-dimensional feature extracted by the feature extraction part 2. It is for giving the special effects in consideration of the three-dimensional features, such as three-dimensional shape for [ in the color picture concerned ] photography, and physical relationship (it adds).

[0099]

It is combining the picture (here, it is also called an added picture) expressed by CG (computer graphics), and, specifically, special effects are added to a color picture. Under the present circumstances, the depth information of the unevenness in a color picture (scene) extracted by the feature extraction part 2, By what kind of object existing in a color picture, or (into what kind of field can it divide?) utilizing the three-dimensional features, such as physical relationship of each object in a color picture, and cubic shape of each object. A context, a collision state, etc. of a virtual body and a scene are distinguished, a virtual body is transformed if needed, and it compounds to a color picture.

[0100]

Here, the case where the virtual body "ball" given as data of 3D(three dimension) CG is compounded is considered by making the color picture (reflected scene) shown in drawing 7 into an example. Although drawing 16 shows the Lord of each object which is a candidate for

photography in the color picture of drawing 7 the physical relationship of a depth direction, as mentioned above, the three-dimensional feature in such a scene is acquired from the feature extraction part 2.

[0101]

Now, a "ball" considers the special effects of moving the place of depth position C in drawing 16 to the left from the screen right. In this case, since the three-dimensional position on which the "ball" which is a virtual body (virtual object) is put, and its configuration information are known, it is comparing the three-dimensional feature (depth value in each position of a scene) of a scene with this, The context of a "ball" and the position of each object in a scene can be distinguished.

[0102]

As shown in drawing 17, a ball passes along a background (depth position D in drawing 16) front, but it is more possible than this to compound a "ball" so that it may pass along a person's (depth position B) back. Thus, it is possible to add a virtual body to a color picture in three dimensions based on the feature extracted by the feature extraction part 2.

[0103]

If similarly the picture of a virtual body "ball" is combined so that a "ball" may move the place of depth position A in a color picture to the left from the right, a "ball" will serve as special effects which pass along a person front through the back of a can.

[0104]

This can be promoted and the collision judgement of a virtual body and the object in a color picture can also be performed. The position of the object in a color picture can be pinpointed in three dimensions from a depth direction an image plane top so that clearly also from drawing 16. Therefore, precise special effects can also be attached for the three-dimensional position pinpointed about each of each object in a color picture by a basis. One of them is special effects of a "collision."

[0105]

Now, the special effects which the virtual body "ball" moves to the place of depth position B in drawing 16 from screen right-hand side are considered. Under the present circumstances, since it turns out that the portion in which a person is is depth position B according to the feature extracted by the feature extraction part 2, when a "ball" comes to a person's position, it turns out that it collides with a person. Then, as shown in drawing 18, the special effects that a "ball" collides with a person and rebounds can also be added. Under the present circumstances, it is possible to also calculate the direction which rallies after colliding by seeing objective three-dimensional shape. it is more effective, if the special effects of expressing the "star" which is one of the effect expressions of a "collision" are added to the collision place (drawing 18 portion of a person's cheek) of the object in a color picture like drawing 18 when colliding.

[0106]

The following special effects are also possible based on the three-dimensional position pinpointed about each of each object in a color picture. For example, the case of the special effects by which a virtual body runs in three dimensions is explained. The example of the special effects which two virtual bodies "ball" move to the depth direction side of the picture concerned from color picture this side is shown in drawing 19. As shown in drawing 19, according to the feature extracted by the feature extraction part 2, it turns out that a can is in depth direction this side, and a person is in the back. As this information is combined with a motion of a "ball" and it was shown in drawing 19, one of two "balls" can hit and rebound upon a nearby can, and another can add the special effects of asking a long distance person and rebounding.

[0107]

As shown in drawing 20, it is also possible to add special effects which turn the virtual body "ball" around the surroundings of a person's head, or to add special effects which turn a "ball" around the surroundings of a finger, as shown in drawing 21.

[0108]

By using the three-dimensional feature of the three-dimensional shape of the object in a color picture, i.e., unevenness, as shown in drawing 22, the special effects of hanging down a fluid with the viscosity of paint etc., etc. from on a person can also be applied. Since the portion of a person's nose has the features, such as protruding rather than other faces, like drawing 22, it is the feature that that avoid the portion of a nose and paint flows can reappear.

[0109]

Since the feature of the three-dimensional physical relationship of the object in a color picture is also known, paint can control the motion and shape of a virtual body of paint not to flow into the portion of the finger of a right arm, and the portion of a can.

[0110]

As it said sitting the "character" of imagination on the portions of a can top or human being's shoulder, or the "dragonfly" of imagination stopping at the portion on a finger, it is also possible to combine various kinds of pictures in the position of a request of the depth direction in a color picture.

[0111]

By what the "ball" of imagination is in middle of the screen, and a hand is moved for to it as shown in drawing 23. According to a motion of the hand concerned, the special effects of giving a motion to a virtual thing (in this case, a "ball" is struck by hand and flown) can also be given to a "ball" in video, using the information on the three-dimensional features, such as a motion of a hand which changes in time, so that it may say that a "ball" moves.

[0112]

Special effects are not only this although the special effects which combine the picture of a virtual body to a color picture were explained as an example above. For example, the special effects by picture composition of erasing background parts, and substituting for another background, or transposing to the background created by CG are also possible.

[0113]

The special effects of leaving as it is and making only background parts color black and white and sepia etc. are also possible for the object (at the example of drawing 7, they are the portions of a person and a can) in a foreground. On the contrary, it can also be said that a mosaic is applied only to a person portion etc. Special effects may be carried out by performing modification of all or some of color pictures. For example, the special effects of swelling the portion of a person's face like a balloon, or wilting it are also possible.

[0114]

As explained above, in the processing section 3, various special effects in consideration of a three-dimensional element are added to the color picture acquired as a natural optical image by the image acquiring part 1 based on the three-dimensional feature extracted by the feature extraction part 2.

[0115]

As an example of special effects, when compounding a virtual body to a color picture, for example, Since the position of the depth direction in the color picture concerned can be defined, the picture of the virtual body concerned is combined to the color picture concerned so that the

virtual body concerned may actually exist in the position of the depth direction in three-dimensional space. It is the same even if it is a case where a color picture is an animation even if it is a case where the picture of a virtual body is an animation, in that case.

[0116]

Since the collision of a motion of the object in a color picture, a position and a motion of a virtual body, the object in [ a position to ] a color picture, and a virtual body can be judged, A motion of the virtual body corresponding to this collision can be expressed, or the display showing the collision having occurred can be performed (for example, the picture of a virtual body showing a collision like a "star" is combined).

[0117]

When compounding a virtual body to a color picture, Since a motion and shape of a virtual body are controllable according to the three-dimensional shape of the object in the color picture concerned, i.e., unevenness, the picture of the virtual body concerned is combined to the color picture concerned so that the virtual body concerned may actually exist in three-dimensional space. It is the same even if it is a case where a color picture is an animation even if it is a case where the picture of a virtual body is an animation, in that case.

[0118]

moreover -- since it is extracted as a feature with three-dimensional object and background parts for photography in a color picture -- making the object and background parts for photography color black and white, sepia, etc. separately, respectively, applying a mosaic, or changing for example, \*\*\*\* -- etc. -- what was said can be performed.

[0119]

How to use the feature explained by this embodiment and special effects are examples to the last, and are not limited to this.

[0120]

Here, for convenience, although explained using virtual bodies, such as a "ball", it is not limited to this. Various things, such as a character, a vehicle, and a building, can be considered as a virtual body. As for a virtual body, what was used as the data of 3DCG using the actual photograph, and the thing which used as CG data a mark, a character, etc. of the handwriting inputted by the user are also included. Although this embodiment explained as what a virtual body does not transform, it is not a thing limited to this and you may change freely according to the kind of virtual body. The picture of a virtual body may be an animation.

[0121]

An input part for a user to input a handwritten mark and character, a vehicle, a building, and the other pictures of various things may newly be added to the composition shown in drawing 1. And CG-data-izing or the treating part for forming 3D CG data is also needed so that the handwritten picture inputted through this input part can be processed by the above-mentioned processing section 3.

[0122]

As explained above, in combining the picture of a virtual body in the processing section part 3 to a depth color picture (it may be a color picture which does not include depth information), The picture of the virtual body concerned is combined to the depth color picture concerned based on at least one of the three-dimensional physical relationship of the for concerned a virtual body and for photography when the position of the virtual body concerned is provided in the depth direction in the depth color picture concerned, and the three-dimensional shape for photography.

[0123]



For example, it controls so that a motion of a virtual body suits the three-dimensional shape of a photography object, and the picture of the virtual body concerned is combined to the depth color picture concerned.

[0124]

The position of the depth direction where the candidate for photography in a depth color picture exists by the special extraction part 2, and the background region of the back for [ concerned ] photography are extracted, and in the processing section 3, the picture of the virtual body concerned is combined to the depth color picture concerned so that a virtual body may exist between the background regions for photography.

[0125]

The 1st position of the depth direction where the 1st candidate for photography that is one of for [ in a depth color picture / two or more ] photography in the special extraction part 2 exists, Extract the 2nd position of the depth direction where the 2nd candidate for photography that is one of everything [ the ] but the inside for [ in the back 1st for / concerned / photography / above-mentioned / two or more ] photography exists, and in the processing section 3. The picture of the virtual body concerned is combined to the depth color picture concerned so that a virtual body may exist between the 1st position concerned and the 2nd position.

[0126]

By the feature extraction part 2, extract the 3rd position as a three-dimensional position in which the candidate for photography in a depth color picture exists, and in the processing section 3.

When the virtual body concerned results in the 3rd position of the above based on a motion of a virtual body and the motion for photography, it judges that the virtual body concerned collided with the candidate for photography concerned, and it controls so that a motion and expression of the object concerned correspond to a collision, and the picture of the virtual body concerned is combined to the depth color picture concerned. When it judges that the virtual body collided with the candidate for photography, effect expressions (for example, "star" etc.) of a collision may be further compounded to the depth color picture concerned.

[0127]

By the feature extraction part 2, extract the 3rd position as a three-dimensional position in which the candidate for photography in a depth color picture exists, and in the processing section 3.

Based on the 3rd position of the above for [ concerned ] photography, and the 4th position of being a three-dimensional position of the virtual body in the above-mentioned depth color picture, a motion and expression of the virtual body concerned are controlled and the picture of the virtual body concerned is combined to the depth color picture concerned.

[0128]

By the feature extraction part 2, each picture for [ two or more ] photography and the picture of the background region after them are extracted from a depth color picture, and at least one of the picture of the above-mentioned background region and each pictures for [ two or more ] photography is processed by the processing section 3. making the object and background parts for example, for photography color black and white, sepia, etc. separately, respectively, applying a mosaic, or changing \*\*\*\* -- etc. -- it carries out.

[0129]

<Picture presentation part>

Next, the picture presentation part 4 is explained.

[0130]

The picture presentation part 4 is for showing a user the color picture (it is hereafter called a

picture with special effects) in which special effects were given by the processing section 3, and the picture with special effects received in the communications department 5.

[0131]

Specifically, the picture presentation part 4 displays on a display the picture with special effects which comprised a display device and was generated by the processing section 3. The picture with special effects received in the communications department 5 is displayed on a display.

[0132]

As shown in drawing 24, it is also possible to divide into the area which displays the picture with special effects generated by the processing section 3 in the display area on the display screen of a display device, and the area which displays the picture with special effects received in the communications department 5, and to show both sides simultaneously. In drawing 24, the picture with special effects received in the communications department 5 is displayed by the area A1, for example, and the picture with special effects generated by the processing section 3 is displayed on the area A2, for example.

[0133]

By using depth information, the picture presentation part 4 can three-dimensional-model-ize the picture with special effects generated by the processing section 3, and can also show it as a scene of 3D (three dimension). If it is made the scene of 3D, it will also become possible to change the position of a viewpoint, and to see or to carry out a corporal vision.

[0134]

<Communications department>

Finally, the communications department 5 is explained.

[0135]

The communications department 5 transmits the picture with special effects generated by the processing section 3 to other terminal units, or receives the same picture with special effects as the above transmitted from other terminal units (for example, it has the same composition as drawing 1).

[0136]

The communications department 5 may be based on the case where it is based on the means of communication of a cable, and the means of communication of radio. First, the case where it is based on the means of communication of radio is explained.

[0137]

For example, in this case, PDC (Personal digital cellular) and CDMA (Code-Division Multiple Access), It communicates with external instruments, such as other terminal units, using a wireless communication system which is used for a cellular phone called PHS. Thereby, transmission of the picture with special effects to an external instrument and reception of the picture with special effects from an external instrument are performed. The means of communication can use the wireless communication system of the wireless LAN which is not limited to portable telephone communication and specified to IEEE802.11a/b/g etc., Bluetooth (trademark) and infrared ray communication, RF communication, and others.

[0138]

Next, the communications department 5 explains the case where it is based on the means of communication of a cable.

[0139]

In this case, the communications department 5 possesses the interface of USB, IEEE1394, etc., are these methods and communicates with the connected external instrument. And transmission

of the picture with special effects to an external instrument and reception of the picture with special effects from an external instrument are performed. transmitting a picture with special effects to connected PC by USB \*\*\*\* -- etc. -- it is . It is not limited to this and the means of communication can use the method of serial communication, a common telephone network, an optical fiber, and others. Transmission and reception of the picture with special effects to another external instrument may be performed via the means of communication which the external instrument has. This can consider performing the transmission and reception to another apparatus on the Internet, for example via the Internet connectivity function of PC (personal computer) by which USB connection is made.

[0140]

<Communication between terminal units>

In [ using the terminal unit in this embodiment described above two or more ] each of these terminal units, As mentioned above in the color picture acquired in real time, special effects are given and a picture with special effects is generated, and it becomes possible using it to communicate among two or more terminal units concerned.

[0141]

Then, it explains, giving some examples about the communications system realized using the terminal unit of this embodiment.

[0142]

Now, the cellular phone with a camera function has spread widely. This is attaching the function photoing and enjoying a photograph and video in addition to the function of the conventional cellular phone, and a photography thing to e-mail etc., and sending them, A partner's figure which transmitted to real time and was simultaneously moved with the partner's cellular phone with a camera function against its figure by receiving in real time. [ who moved using the function and camera which communicate with others ] The function etc. of the TV phone which communicates while looking at both pictures are put in practical use.

[0143]

The communications system of this embodiment is realized in the form where this is replaced, for example. It is an image of the new portable telephone system which can perform picture communication which was explained in the top.

[0144]

For example, a user photos his figure and scenery, the worried thing, etc. using the cellular phone as a terminal unit concerning this embodiment which constitutes this new portable telephone system. And a user acquires the picture with special effects (of course, there may be video in this picture) generated using the depth information corresponding to that picture. Using this, he enjoys himself to the default window of a cellular phone, etc. This picture (or video) is attached to e-mail, and it sends to a partner, and uses as a means of communication. During the real time picture communication of a TV phone, it can tell well against its feeling, or in order to give entertainment nature to communication, special effects can be added like the above-mentioned.

[0145]

Then, with reference to the flow chart which shows drawing 33 and drawing 34 a series of flows using the example of a TV phone, it explains concretely. Now, the two persons A and B hold the cellular phone with which both sides have the composition shown in drawing 1. The persons A and B are photoing their face mutually. In the state where no special effects are added, the persons' A and B cellular phone is mutually connected via the communications department 5 of mutual apparatus, the picture of the person's A face is shown through the person's B picture

presentation part 4, and the person A is shown the picture of the person's B face conversely. The TV phone using a picture is realized by this. And at a certain time, the person A thinks up that he will attach the special effects that \*\*\*\* is flying to the surroundings of its face, and photos the picture of his face with the cellular phone concerned. Thereby, the image acquiring part 1 acquires a picture (namely, depth color picture) including picture depth information (Step S1). In the feature extraction part 2, the three-dimensional feature in the face picture concerned is extracted based on depth information etc. (Step S2). If the person A chooses the picture to which \*\*\*\* is flying, in the processing section 3, the picture chosen as the butterfly was flying around the face about will be combined based on the three-dimensional feature extracted by the face picture concerned, and a picture with special effects will be generated (Step S3). While displaying this picture with special effects on the picture presentation part 4 (step S4), the picture with special effects concerned is transmitted to the person B with the person's A transmission instruction (Step S5). In the cellular phone which the person B possesses, since it will be displayed on the picture presentation part 4 if the picture with special effects concerned is received in the communications department 5 (Step S6) (Step S7), the person B can see the face of the person A with the special effects. Thus, it becomes possible to add the entertainment nature which was not the conventional TV phone to communication.

[0146]

This embodiment is not only a case like a cellular phone. Next, another example is explained. From the former, the camera of USB connection is connected to a personal computer (PC), and the scene etc. are always photoed. And connecting the PC with the Internet and exhibiting the picture of the real time currently photoed generally is performed. This is usually service called a live camera or a fixed point observation camera in many cases. What is photoed is going across the situation etc. of the state of the situation of its room, a pet's appearance, and the entering condition of the visitor of a store who is managing, and the rows of houses of Shibuya variably variously.

[0147]

In such directions for use, special effects can be added to such a picture by connecting the terminal unit concerning this embodiment to PC instead of the above-mentioned USB connection camera. Therefore, for example, the store which sells a certain goods lets a live camera pass, and presupposes that the picture of goods is opened to the public on the Internet now. And it is assumed that some goods are photoed, changing direction of a camera (pan operation). Under the present circumstances, when goods with what is called "high recommendation" are photoed, around that product, When the special effects which turn the virtual body (object) showing "high recommendation" like drawing 20 were applied and the goods of a new product were photoed, it was able to be said that special effects around which the virtual body (object) showing the new product concerned turns were applied. Thus, entertainment nature can be given to the picture of a live camera etc. and additional information, such as "high recommendation" and a "new product", can be provided further.

[0148]

(The 1st modification of a 1st embodiment)

Drawing 25 shows the example of composition of the important section of the terminal unit concerning the 1st modification.

[0149]

The difference between the terminal unit shown in drawing 25, and the terminal unit shown in drawing 1, While the feature extraction part 2 extracts the three-dimensional feature for

photography like the above-mentioned in drawing 25 based on the depth information included in the depth color picture acquired by the image acquiring part 1, The three-dimensional feature for photography is extracted also from the depth color picture received in the communications department 5 based on the depth information included there.

[0150]

The processing section 3 of drawing 25 is constituted so that special effects may be given to a picture with special effects including the depth color picture (or color picture) received in the image acquiring part 1 or the communications department 5, and depth information like the above-mentioned based on the feature extracted by the feature extraction part 2.

[0151]

Here, since the picture with special effects is generated from the color picture in which depth information is matched from the first, or the depth color picture, depth information is matched also with the picture with special effects (contained).

[0152]

It becomes possible to give special effects by having such composition also about the depth color picture and the picture with special effects which were received not only in the depth color picture obtained by the image acquiring part 1 but in the communications department 5.

[0153]

Next, operation of the terminal unit which has the composition shown in drawing 25 is concretely explained using the example of a previous cellular phone, referring to the flow chart shown in drawing 35.

[0154]

Now, two persons hold the cellular phone (it had a function concerning the 1st modification) of composition as both sides showed drawing 25, and it is assumed that communication using a TV phone is performed between two persons. Under the present circumstances, although it was possible in a 1st embodiment to have applied special effects to the picture which he is photoing, In the 1st modification, in the communications department 5. While receiving a picture including the depth information sent by the partner (Step S11) and displaying this (Step S12), the three-dimensional feature of the received picture concerned is extracted based on the depth information etc. of the received picture concerned (Step S13). Therefore, in the processing section 3 of a receiver, it becomes possible to add special effects also to the received picture concerned (Step S14). The thing which show the picture presentation part 4 and which is both transmitted to a partner further again can also do the picture with special effects which added desired special effects and was generated (Step S16).

[0155]

Therefore, while both sides move their face and are, for example, performing communication using a TV phone, a trick can be played on a partner's face. For example, in the conte program of television, as expression and the punishment game of anger, a "tub" may be dropped on a person and laughter may often be taken. Like this, in communication with a partner, when something feels unpleasant, a partner's face picture is received, The new picture communication which gives special effects to which a "tub" is dropped, and he looks at and enjoys of urging cautions further while sending against it and delighting is attained.

[0156]

(The 2nd modification of a 1st embodiment)

Drawing 26 shows the example of composition of the important section of the terminal unit concerning the 2nd modification.

[0157]

The difference between the terminal unit shown in drawing 26, and the terminal unit shown in drawing 1, In drawing 26, it has the composition that the added picture storage parts store 7 which memorizes two or more added picture data (for example, CG data and 3D CG data) of a virtual body etc. further, and the special-effects selecting part 6 for choosing a desired added picture and special effects out of this were added by the composition shown in drawing 1.

[0158]

The special-effects selecting part 6 shows a user the kind of an addable added picture or special effects with reference to the added picture memorized by the added picture storage parts store 7. A user chooses an added picture or special effects to add out of what was shown, and this information is passed to the processing section 3. The processing section 3 adds the special effects according to the kind of the added picture which the user chose, or special effects. In the special-effects selecting part 6, a user may not be made to choose the kind of an added picture or special effects, but an added picture and special effects may be chosen at random.

[0159]

By changing in this way, a user chooses the virtual body compounded to the acquired color picture, and special effects, and it becomes possible to generate the color picture with special effects according to a user's liking.

[0160]

(The 3rd modification of a 1st embodiment)

Drawing 27 shows the example of composition of the important section of the terminal unit concerning the 3rd modification. The difference between the terminal unit which transformed the 3rd modification for the 2nd modification of the above further, and was shown in drawing 27, and the terminal unit shown in drawing 26, In drawing 27, it has the composition that the scene analyzing parts 8 which analyze the feature of the scene further reflected in the depth color picture (or color picture) acquired by the image acquiring part 1 in the composition shown in drawing 26 were added.

[0161]

The scene analyzing parts 8 analyze the feature of the scene reflected in the depth color picture (or color picture) acquired by the image acquiring part 1. The feature of a scene refers to specification of what is reflected in the scene, the color atmosphere of a scene, the composition of a scene, etc. For example, if the object reflected in the scene is a person, it will distinguish that he is a person. And the information is given to the special-effects selecting part 6. In the special-effects selecting part 6, adding to a person limits the kind of special effects to a suitable thing. When what is reflected in the scene is longwise, and the picture of the virtual body (object) which revolves around it horizontally is added (for example, composition), it is effective, but. When very oblong, it may be better to add the picture of the virtual body (object) which revolves around it perpendicularly, and there is that it is also more effective to choose the special effects of combining the picture of another virtual body which is not a virtual body with the rotating motion. There is that it is also better to change the shape and the colors of an added picture, such as a virtual body, variously from the overall tone of a scene. For example, probably, there are few effects, even if that of red taro compounds a red virtual body on the scene reflected mostly on the whole. It may be made to change a motion of a virtual body according to the composition of the picture (scene) itself, such as arrangement etc. of the object which has moved to the scene.

[0162]

By composition shown in drawing 27 by as mentioned above, the thing for which the feature of a

scene is analyzed by the scene analyzing parts 8. According to the feature of the scene concerned, parameters, such as a motion of the virtual body etc. to add, a color, and shape, etc. can be changed, or the kind of special effects corresponding to the feature of the scene concerned can be chosen, and special effects with a higher effect can be added.

[0163]

(The 4th modification of a 1st embodiment)

The composition of the terminal unit which showed drawing 1 the terminal unit concerning the 4th modification, the composition of the 1st modification shown in drawing 25, It has the composition that the selecting part 9 which can choose further the picture which communicates in the communications department 5 as either the composition of the 2nd modification shown in drawing 26 or composition of the 3rd modification shown in drawing 27 was added.

[0164]

Drawing 28 shows the example of composition of the important section of the terminal unit concerning the 4th modification, and shows the terminal unit which adds the selecting part 9 to the composition of the 1st modification shown in drawing 25, and is constituted by drawing 28.

[0165]

As shown in drawing 28, it becomes possible by adding the selecting part 9 to choose the data exchanged for other terminal units via the communications department 5.

[0166]

The effect acquired by the 4th modification is explained using the example which the 1st modification used by the way. The 1st modification and the same example are considered. The cellular phone with the function of composition as two persons showed drawing 28 which both sides require for the 4th modification now is held, and it is assumed that communication using a TV phone is performed between two persons. And both sides move their face and are performing communication using a TV phone. Although communication using various special effects is performed to this midst, only the special effects which play a trick on a partner's face perform selection of not transmitting to a partner etc. Thereby, although the picture with special effects is always usually sent to the partner, when something feels unpleasant, special effects to which a "tub" is dropped are given to a partner's face picture, and it becomes possible only of the picture with special effects for him to only see and enjoy himself and not to send to a partner.

[0167]

Not only in communication of 1 to 1 but in the communication (one pair in large numbers in large numbers opposite large number) with a lot of people, an effect comes out of this further. For example, it is assumed that communication using the TV phone machine which has the composition which a certain person D showed simultaneously to the three persons A, B, and C and drawing 28 was performed. At a certain time, the person D felt unpleasant to the person's A speech and conduct, and made selection which gives the special effects to which a "tub" is dropped. However, since it thought that it was quarreling if it was sent to A, it is sent to the persons B and C and selection of the transmission object of having conveyed one's feeling etc. is attained.

[0168]

Although the selection to transmission was explained, it chooses similarly, and it is made not to receive the picture with special effects from a specific person, or may be made to make it not receive above the picture with special effects to which specific special effects were given to reception.

[0169]

(The 5th modification of a 1st embodiment)

It is also possible to add a sound effect suitable for the picture concerned to a color picture with special effects.

[0170]

The sound effect adjunct for adding a sound effect suitable for the picture concerned to a color picture with special effects can also be further added to the composition shown in drawing 1, the composition of the 1st modification shown in drawing 25, the composition of the 2nd modification shown in drawing 26, or composition of the 3rd modification shown in drawing 27.

[0171]

Thereby, a sound effect can be attached to a color picture with special effects, and special effects can be heightened further. For example, although the virtual body asked the person, and rebounded and the example of the special effects which combine the picture of virtual bodies, such as a "star" etc. showing having collided on that occasion, was explained by a 1st above-mentioned embodiment, in the composition shown in drawing 1, addition of the above-mentioned sound effect adjunct (not shown) will add an impact noise in the case of this collision -- things can be carried out and effect of the special effects concerned can be achieved.

[0172]

As explained above, according to a 1st embodiment of the above, with the picture for photography (for example, here natural optical image). (The catoptric light picture which includes depth information in each pixel is acquired) Since the information on the depth direction of the natural optical image concerned (depth information) is acquirable, The three-dimensional features, such as three-dimensional shape for [ in a natural optical image ] photography and three-dimensional physical relationship for photography, are extracted using this depth information, Based on this three-dimensional feature, the picture (added picture) expressing effect expressions and virtual bodies, such as a shock, can be combined, and the special effects which employed the three-dimensional feature for [ in a picture ] photography efficiently can be added.

[0173]

(A 2nd embodiment)

Next, a 2nd embodiment of this invention is described.

[0174]

<The whole composition>

Drawing 29 shows the outline composition of the whole communications system concerning a 2nd embodiment. As shown in drawing 29, this communications system, Two or more terminal units which are radio communications systems, for example, have the composition shown in drawing 1 (for example, here.) the -- one -- a terminal -- 201 -- the -- two -- a terminal -- 202 -- two -- a \*\* -- the -- one -- a terminal -- 201 -- wireless connection -- carrying out -- a base station device -- (-- BS --) -- 211 -- the -- two -- a terminal -- 202 -- wireless connection -- carrying out -- a base station device -- (-- BS --) -- 212. It comprises a predetermined communications network (network) which connects the base station device 211,212 and the server apparatus 200. Via the base stations 211 and 212 and a network, the server apparatus 200 of each other is connected with the 1st terminal 201 and the 2nd terminal 202 so that communication is possible.

[0175]

As the 1st terminal 201 and 2nd terminal 202 were mentioned above, picture communication can be performed by transmitting and receiving mutually a picture with special effects, and the picture which does not attach special effects. Since the 1st terminal 201 and 2nd terminal 202



have the composition shown in drawing 1, respectively, At each terminal, the depth information corresponding to the color picture concerned, i.e., the depth color picture as a color picture which includes depth information here, is acquirable besides a color picture.

[0176]

Between the 1st terminal 201 and the 2nd terminal 202, when performing picture communication, from the communications department 5 of each terminal, the depth information corresponding to the picture concerned is also transmitted with the picture which does not attach a picture with special effects, or special effects. Here, depth information shall be included in the picture itself which transmits like the above-mentioned.

[0177]

Between the 1st terminal 201 and the 2nd terminal 202, when transmitting and receiving a picture with special effects, from the communications department 5 of each terminal, the three-dimensional feature extracted by the feature extraction part 2 of each terminal with the picture with special effects including the depth information concerned is also transmitted.

[0178]

The server apparatus 200 between the 1st terminal 201 and the 2nd terminal 202, It is installed and managed by the provider who provides the service which can perform picture communication, When performing the picture communication concerned between the 1st terminal 201 and the 2nd terminal 202, It is received by the server apparatus 200 and the picture transmitted from the 1st terminal 201 and 2nd terminal 202, depth information, etc. are certainly transmitted to a mating terminal via this server apparatus 200.

[0179]

Now, in this 2nd embodiment, the feature is that it compounds an advertisement image based on the depth information corresponding to the picture concerned in the above-mentioned server apparatus 200 at an added picture, i.e., here, at the picture transmitted from the 1st terminal 201 and 2nd terminal 202.

[0180]

Drawing 30 is what showed the example of composition of the server apparatus 200, and comprises the receive section 10, the feature extraction part 2, the processing section 3, and the transmission section 11. In drawing 30, identical codes are given to drawing 1 and identical parts, and a different portion is explained. Namely, the server apparatus 200 shown in drawing 30, It replaces with each of the image acquiring part 1 of the terminal unit of drawing 1, the picture presentation part 4, and the communications department 5, By the receive section 10 which receives the picture (picture which does not attach a picture with special effects including depth information, or special effects) transmitted from the 1st terminal 201 and 2nd terminal 202, and the processing section 3 of the server apparatus 200. The transmission section 11 which transmits the picture with an advertisement generated by giving special effects, such as composition of an advertisement image, to the picture received in the receive section 10 to the 1st terminal 201 or 2nd terminal 202 that is the address concerned of a picture which received is provided.

[0181]

Next, the processing operation of the server apparatus 200 is explained with reference to the flow chart shown in drawing 36.

[0182]

If the receive section 10 of the server apparatus 200 receives a picture (picture which does not attach a picture with special effects, or special effects) including the depth information

transmitted from the 1st terminal 201 and 2nd terminal 202 (Step S21), The feature extraction part 2 of the server apparatus 200 extracts the three-dimensional feature in the picture concerned from the depth information corresponding to the picture received in the receive section 10, and the picture concerned like the feature extraction part 2 of the terminal unit of drawing 1 (Step S22).

[0183]

In the processing section 3 of the server apparatus 200, based on the picture received in the receive section 10, and the extracted three-dimensional feature concerned, an advertisement image is compounded and special effects are given to the received picture concerned (Step S23). The picture with an advertisement generated by compounding an advertisement image by this processing section 3 is transmitted to the 1st terminal 201 or 2nd terminal 202 that is an address at a basis also as the one concerned of a picture which received from the transmission section 11 (Step S24).

[0184]

Here, the processing section 3 of the server apparatus 200 is explained. the picture (the picture with special effects to which special effects are already given at the 1st and 2nd terminal 201,202.) which the processing section 3 received in the receive section 10 Or special effects are given by compounding arbitrary advertisement images further from the general picture to which special effects are not attached in the picture concerned received in the receive section 10 in consideration of the three-dimensional feature acquired by the feature extraction part 2.

[0185]

In the picture concerned, it is compounding further the advertisement expressed by CG (computer graphics) etc., and, specifically, a picture with an advertisement is generated. Under the present circumstances, the three-dimensional shape of unevenness for [ which was extracted by the feature extraction part 2 / in a picture ] photography (object), What kind of object exists in a scene, or (into what kind of field can it divide?) it is made to be the same as that of the processing section 3 of drawing 1 based on the three-dimensional features, such as physical relationship of each object in a scene, and cubic shape of each object, Physical relationship, such as a context etc. of the object in an advertisement image and the received picture, a collision, etc. are distinguished, and an advertisement is compounded.

[0186]

Here, it supposes that the picture shown in drawing 7 was received in the receive section 10, the case where an advertisement (picture) is compounded in the picture (it is hereafter called a reception picture) shown in this drawing 7 is taken for an example, and the processing operation of the processing section 3 of the server apparatus 200 is explained.

[0187]

The additional means of the easiest advertisement compound an advertisement to background parts so that an advertisement may not be applied to the foreground of a reception picture. Drawing 31 shows signs that the advertisement image was compounded to the reception picture of drawing 7. As shown in drawing 14, according to the three-dimensional feature extracted by the feature extraction part 2, a having mentioned above passage can distinguish a foreground and a background easily. An advertisement is added using this three-dimensional information. Even if the person who is a foreground moves by carrying out like this, it becomes possible to always show a background an advertisement so that an advertisement may not be applied to a person. By carrying out like this, it becomes possible to always show an advertisement, without having an adverse effect on communication.

[0188]

It is effective even if it adds an advertisement to the virtual body (compounded) added by the processing section 3 of the terminal unit in a reception picture. For example, it is assumed that the special effects that the virtual body "ball" moved the back of the person who explained to the reception picture by drawing 17 were already given. In the processing section 3 of the server apparatus 200, an advertisement is stuck on this virtual body "ball" itself, or as shown in drawing 32, an advertisement is compounded so that an advertisement may be hung down from a "ball." By carrying out like this, the advertisement itself will move about the inside of a screen in accordance with a motion of a "ball." That is, advertising effectiveness is also included into the entertainment effect by the special effects added by the terminal side included in a reception picture. Thereby, without giving a user a bad impression, while giving the entertainment effect, it becomes possible to carry out advertising presentation nonchalantly.

[0189]

The virtual body itself may serve as an advertisement and it compounds the virtual body united with the advertisement to a reception picture in the processing section 3 of the server apparatus 200. As for the virtual body united with the advertisement, things, such as a CG image of newly released goods, are mentioned, for example.

[0190]

The additional means of the advertisement explained above are examples to the last, and are not limited to this. An advertisement can be freely added using the three-dimensional feature extracted by the feature extraction part 2 of the server apparatus 200. For example, although the example of the special effects from which paint falls to the head was shown in drawing 22, the character in which the package of a certain goods instead of paint could fall, and this showed a trade name, a release date, etc. may fall. Although the collision judgement was considered as the virtual body and the example which attaches virtual bodies, such as a star, in that case was shown in drawing 18, an advertisement may be included and shown into this star.

[0191]

It is also possible to place two or more another advertisements in a screen, and when a reception picture is an animation, the advertising contents may change by a scene. for example, if the person reflected to the scene is a woman, and it is a male about a female-oriented advertisement, as it said that a male-oriented advertisement was added, it will come out.

[0192]

In the 1st terminal 201 and 2nd terminal 202. Since the feature three-dimensional at each feature extraction part 2 is extracted from the first, If this extracted three-dimensional characteristic information is also certainly transmitted with a picture (picture which does not attach a picture with special effects including depth information, or special effects), it will become unnecessary to have the feature extraction part 2 in the server apparatus 200. And as shown in drawing 37, an advertisement image is compoundable [ based on the characteristic information corresponding to the reception picture concerned similarly received in the receive section 10 ] in the server apparatus 200, like the above-mentioned to the picture (reception picture) received in the receive section 10 (Step S31 - Step S32).

[0193]

According to the communications system using the server apparatus 200 concerning a 2nd embodiment, it becomes possible to provide the picture which added the advertisement effectively into the picture to transmit in addition to picture communication which was explained by a 1st embodiment. By this showing the position which does not become the obstacle of

communication between users an advertisement, or already incorporating an advertisement into a certain virtual body, while giving entertainment to an advertisement, it becomes possible to show a user an advertisement, without having an adverse effect.

[0194]

It combines suitably and each above embodiment and its modification can be carried out.

[0195]

It is also possible to realize processing in the embodiment of the invention in this application by the program which can be executed by computer, and to realize this program by computer as a storage which can be read. For example, a part of image acquiring part 1 among the composition of the terminal unit of drawing 1 (for example, the natural optical image image pick-up part 103, the catoptric light picture image pick-up part 108, and the imaging operation control section 104). Since being stored in one integrated circuit is desirable as mentioned above, the other image acquiring part 1 and the feature extraction part 2, and the processing section 3 at least can also be realized as software. It is realizable by installing in the terminal unit concerned the program for making the processing operation shown in drawing 33 - drawing 35 of the terminal unit concerning the above-mentioned embodiment perform.

[0196]

As a storage, a magnetic disk, a floppy disk, a hard disk, Optical discs (CD-ROM, CD-R, DVD, etc.), magneto-optical discs (MO etc.), semiconductor memory, etc. can memorize a program, and as long as it is a storage which a computer or an embedded system can read, the memory form may be which gestalt.

[0197]

OS (operation system) which is working on a computer based on directions of the program installed in the computer or the embedded system from the storage, A part of each processing for MW(s) (middleware), such as database management software and a network, etc. to realize this embodiment may be performed.

[0198]

The storage which the above-mentioned storage downloaded the program transmitted by not only the medium that became independent of a computer or an embedded system but LAN, the Internet, etc., and was memorized or stored temporarily is also contained.

[0199]

Also when a storage is not restricted to one but processing in this embodiment is performed from two or more media, it may be contained in the storage in this invention, and the composition of a medium may be which composition.

[0200]

The computer or embedded system in the invention in this application, Based on the program memorized by the storage, it may be for performing each processing in this embodiment, and the device which consists of one, such as a personal computer and a microcomputer, and two or more devices may be which composition, such as a system by which network connection was carried out.

[0201]

The apparatus which can realize the function of the invention in this application by a program, and a device are named generically the computer in the invention in this application not only including a personal computer but including an arithmetic processing unit, a microcomputer, etc. which are contained in an information management system.

[0202]

this invention is not limited to the above 1st - a 2nd embodiment, and in the range which does not deviate from the gist, many things are boiled and it can be changed at an execution phase. The invention of various stages is included in the above-mentioned embodiment, and various inventions may be extracted by the proper combination in two or more constituent features indicated. For example, even if some constituent features are deleted from all the constituent features shown in an embodiment, The technical problem (at least one) described in the column of Object of the Invention is solvable, and when the effect (at least one) described in the column of the effect of the invention is acquired, the composition from which these constituent features were deleted may be extracted as an invention.

[0203]

[Effect of the Invention]

As explained above, according to this invention, it is processible to combine an advertisement image and the picture of a desired virtual body etc. based on the three-dimensional features, such as each three-dimensional shape for [ in the photoed picture (video and a still picture are included) / two or more ] photography, and physical relationship.

[Brief Description of the Drawings]

[Drawing 1] The figure showing roughly the example of composition of the important section of the terminal unit concerning a 1st embodiment of this invention.

[Drawing 2] The figure for explaining a usual monochrome image and depth monochrome image.

[Drawing 3] The figure showing the example of composition of an image acquiring part roughly.

[Drawing 4] The figure showing the example of composition of a catoptric light picture image pick-up part.

[Drawing 5] The figure showing the three-dimensional image of the hand in the catoptric light picture concerned produced from the depth information included in the catoptric light picture photoed in the catoptric light picture image pick-up part by making human being's hand applicable to photography.

[Drawing 6] The figure showing one example of a catoptric light picture ((a) figure) and a natural optical image ((b) figure) acquired by the image acquiring part.

[Drawing 7] The figure showing an example of the depth color picture which is a processing object in a feature extraction part with a monochrome image.

[Drawing 8] The figure showing roughly the three-dimensional physical relationship of the object in the depth color picture of drawing 7.

[Drawing 9] The figure showing the background parts in the depth color picture of drawing 7.

[Drawing 10] The figure showing the person portion in the depth color picture of drawing 7.

[Drawing 11] The figure showing the arm part in the depth color picture of drawing 7.

[Drawing 12] The figure showing the can portion in the depth color picture of drawing 7.

[Drawing 13] The figure showing one relation of the depth direction of the object in the depth color picture of drawing 7.

[Drawing 14] The figure showing the physical relationship, and the foreground and the background of a depth direction in the depth color picture of drawing 7.

[Drawing 15] The figure for explaining how to search for the detailed physical relationship of the object in the picture concerned from the break point (jump edge) of the depth direction in the depth color picture of drawing 7.

[Drawing 16] The figure for explaining how to pinpoint the position of the depth direction of each object in the depth color picture of drawing 7.

[Drawing 17] The figure for explaining an example of the picture with special effects which

showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 18]The figure for explaining other examples of the picture with special effects which showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 19]The figure for explaining the example of further others of the picture with special effects which showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 20]The figure for explaining the example of further others of the picture with special effects which showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 21]The figure for explaining the example of further others of the picture with special effects which showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 22]The figure for explaining the example of further others of the picture with special effects which showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 23]The figure for explaining the example of further others of the picture with special effects which showed the result of having added special effects to the depth color picture of drawing 7.

[Drawing 24]The figure showing the example of the screen separation for displaying a picture on a screen presentation part.

[Drawing 25]The figure showing roughly the example of composition of the important section of the terminal unit concerning the 1st modification of a 1st embodiment of this invention.

[Drawing 26]The figure showing roughly the example of composition of the important section of the terminal unit concerning the 2nd modification of a 1st embodiment of this invention.

[Drawing 27]The figure showing roughly the example of composition of the important section of the terminal unit concerning the 3rd modification of a 1st embodiment of this invention.

[Drawing 28]The figure showing roughly the example of composition of the important section of the terminal unit concerning the 4th modification of a 1st embodiment of this invention.

[Drawing 29]The figure showing roughly the composition of the whole communications system concerning a 2nd embodiment of this invention.

[Drawing 30]The figure showing roughly the example of composition of the important section of a server apparatus.

[Drawing 31]The figure for explaining an example of a picture with an advertisement.

[Drawing 32]The figure for explaining other examples of a picture with an advertisement.

[Drawing 33]The flow chart for explaining the processing operation of a terminal unit.

[Drawing 34]The flow chart for explaining the processing operation of a terminal unit.

[Drawing 35]The flow chart for explaining the processing operation of a terminal unit.

[Drawing 36]The flow chart for explaining the processing operation of a server apparatus.

[Drawing 37]The flow chart for explaining the processing operation of a server apparatus.

[Description of Notations]

1 -- Image acquiring part

2 -- Feature extraction part

3 -- Processing section

4 -- Picture presentation part

5 -- Communications department

6 -- Special-effects selecting part

- 7 -- Added picture storage parts store
- 8 -- Scene analyzing parts
- 9 -- Selecting part
- 10 -- Receive section
- 11 -- Transmission section

---

[Translation done.]

(19) 日本国特許庁(JP)

## (12) 公開特許公報(A)

(11) 特許出願公開番号

特開2004-145448

(P2004-145448A)

(43) 公開日 平成16年5月20日(2004.5.20)

(51) Int.Cl.<sup>7</sup>

F I

テーマコード(参考)

G06T 17/40

G06T 17/40

B

5B050

G06T 1/00

G06T 17/40

C

5B057

G06T 3/00

G06T 1/00

315

5L096

G06T 7/60

G06T 3/00

300

G06T 15/70

G06T 7/60

150S

審査請求 未請求 請求項の数 20 O L (全 40 頁) 最終頁に続く

(21) 出願番号 特願2002-307170 (P2002-307170)

(22) 出願日 平成14年10月22日(2002.10.22)

(特許庁注: 以下のものは登録商標)  
フロッピー

(71) 出願人 000003078

株式会社東芝

東京都港区芝浦一丁目1番1号

(74) 代理人 100058479

弁理士 鈴江 武彦

(74) 代理人 100084618

弁理士 村松 貞男

(74) 代理人 100068814

弁理士 坪井 淳

(74) 代理人 100092196

弁理士 橋本 良郎

(74) 代理人 100091351

弁理士 河野 哲

(74) 代理人 100088683

弁理士 中村 誠

最終頁に続く

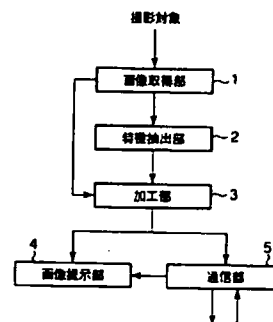
(54) 【発明の名称】 端末装置、サーバ装置および画像加工方法

## (57) 【要約】

【課題】撮影した画像（動画像、静止画像を含む）中の複数の撮影対象のそれぞれの3次元的な形状や3次元的な位置関係などの3次元的な特徴に合わせて、広告画像や所望の画像を合成するなどの加工が行える端末装置を提供する。

【解決手段】複数の撮影対象を含む第1の画像を取得するとともに、1または複数の画素からなる複数の領域から構成される前記第1の画像の前記複数の領域のそれぞれに対応する実行き情報を取得し、前記第1の画像と前記実行き情報とを基に、少なくとも前記複数の撮影対象の3次元的な形状と上記複数の撮影対象の3次元的な位置関係を特徴情報として抽出して、この特徴情報を基に前記第1の画像を加工する。

【選択図】 図1





## 【特許請求の範囲】

## 【請求項1】

互いに通信可能な複数の端末装置のうちの1つである端末装置であって、  
複数の撮影対象を含む第1の画像を取得する第1の取得手段と、  
前記第1の画像は1または複数の画素からなる複数の領域から構成され、この複数の領域のそれぞれに対応する奥行き情報を取得する第2の取得手段と、  
前記第1の画像と前記奥行き情報とを基に、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を特徴情報として抽出する抽出手段と、  
この抽出手段で抽出した特徴情報を基に、前記第1の画像を加工して第2の画像を生成する生成手段と、  
を具備したことを特徴とする端末装置。

10

## 【請求項2】

前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該第1の画像中の奥行き方向に当該物体の位置を定めたときの、当該物体と前記複数の撮影対象の3次元的な位置関係と、前記複数の撮影対象の3次元的な形状とのうちの少なくとも1つを基に、当該物体の画像を当該第1の画像に合成することを特徴とする請求項1記載の端末装置。

## 【請求項3】

前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、前記物体の動きが前記撮影対象の3次元的な形状に合うように制御して、当該物体の画像を当該第1の画像に合成することを特徴とする請求項1記載の端末装置。

20

## 【請求項4】

前記抽出手段は、前記第1の画像中の前記撮影対象の存在する奥行き方向の位置と当該撮影対象の後の背景領域とを抽出し、  
前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該物体が当該撮影対象と前記背景領域の間に存在するように当該物体の画像を当該第1の画像に合成することを特徴とする請求項1記載の端末装置。

## 【請求項5】

前記抽出手段は、前記第1の画像中の前記複数の撮影対象のうちの1つである第1の撮影対象の存在する奥行き方向の第1の位置と、当該第1の撮影対象の後にある、前記複数の撮影対象のうちの他の1つである第2の撮影対象の存在する奥行き方向の第2の位置とを抽出し、  
前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該物体が当該第1の位置と第2の位置の間に存在するように当該物体の画像を当該第1の画像に合成することを特徴とする請求項1記載の端末装置。

30

## 【請求項6】

前記抽出手段は、前記第1の画像中の前記撮影対象の存在する3次元的な位置として第3の位置を抽出し、  
前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該物体の動きと当該撮影対象の動きとに基づき、当該物体が前記第3の位置に至ったとき、当該物体が当該撮影対象と衝突したと判断し、当該物体の動きや表現が衝突に対応するよう制御して当該物体の画像を当該第1の画像に合成することを特徴とする請求項1記載の端末装置。

40

## 【請求項7】

前記生成手段は、前記物体が前記撮影対象と衝突したと判断したとき、衝突の効果表現を前記第1の画像に合成することを特徴とする請求項6記載の端末装置。

## 【請求項8】

50

前記抽出手段は、前記第1の画像中の前記撮影対象の存在する3次元的な位置として第3の位置を抽出し、

前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該撮影対象の前記第3の位置と、前記第1の画像中の当該物体の3次元的な位置である第4の位置とに基づき、当該物体の動きや表現を制御して当該物体の画像を当該第1の画像に合成することを特徴とする請求項1記載の端末装置。

【請求項9】

前記抽出手段は、前記第1の画像から前記複数の撮影対象のそれぞれの画像と、それらの後の背景領域の画像とを抽出し、

前記生成手段は、前記背景領域の画像と前記複数の撮影対象のそれぞれの画像のうちの少なくとも1つを加工することにより前記第2の画像を生成することを特徴とする請求項1記載の端末装置。

【請求項10】

前記第2の取得手段は、

前記第1の画像に対応し、各画素値に奥行き情報を含む第3の画像を取得する手段と、  
前記第3の画像から前記第1の画像の各領域に対応する奥行き情報を抽出する手段と、  
を具備し、

前記抽出手段は、前記第1の画像と当該第1の画像の各領域に対応する前記奥行き情報とから、少なくとも前記撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を抽出することを特徴とする請求項1記載の端末装置。

【請求項11】

前記第2の取得手段は、

前記撮影対象に向けて発光する発光手段と、

この発光手段を発光させて前記撮影対象に照射した光の当該撮影対象からの反射光を含む第1の光量を受光する第1の受光手段と、

前記発光手段が発光していないときに、前記反射光を含まない第2の光量を受光する第2の受光手段と、

前記第1の光量から前記第2の光量を差し引いて、前記第1の光量から前記反射光の成分を抽出することにより、各画素値に奥行き情報を含む第3の画像を生成する手段と、

前記第3の画像から前記第1の画像の各領域に対応する奥行き情報を算出する手段と、  
を具備し、

前記抽出手段は、前記第1の画像と当該第1の画像の各領域に対応する前記奥行き情報とから、少なくとも前記撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を抽出することを特徴とする請求項1記載の端末装置。

【請求項12】

前記第3の画像は前記第1の画像と同時に取得することを特徴とする請求項10または11記載の端末装置。

【請求項13】

前記第1の取得手段で取得する画像はカラー画像であることを特徴とする請求項1記載の端末装置。

【請求項14】

前記生成手段は、前記第1の画像に衝撃などの効果表現や物体を表現した付加画像を合成することにより前記第2の画像を生成するものであって、

この生成手段で前記第1の画像に合成するための複数種類の付加画像を記憶する記憶手段と、

この記憶手段に記憶された複数の付加画像の中から前記第1の画像に合成する付加画像を選択する手段と、

をさらに具備したことを特徴とする請求項1記載の端末装置。

【請求項15】

前記生成手段は、前記第1の画像に衝撃などの効果表現や物体を表現した付加画像を合成

10

20

30

40

50

することにより前記第2の画像を生成するものであって、

この生成手段で前記第1の画像に合成するための複数種類の付加画像を記憶する記憶手段と、

前記第1の画像から、当該第1の画像に写っている物体、色的な雰囲気、画像の構図のうちの少なくとも1つをシーンの特徴として抽出する手段と、

前記シーンの特徴を基に、前記記憶手段に記憶された複数種類の付加画像の中から前記第1の画像に合成する付加画像を選択する手段と、

をさらに具備したことを特徴とする請求項1記載の端末装置。

【請求項16】

前記第2の画像に適した効果音を付加する効果音付加手段をさらに具備したことを特徴とする請求項1記載の端末装置。 10

【請求項17】

前記第2の画像を表示する表示手段と、

少なくとも前記第2の画像を前記複数の端末装置のうちの他の端末装置との間で送受信する通信手段と、

をさらに具備したことを特徴とする請求項1記載の端末装置。

【請求項18】

互いに通信可能な複数の端末装置と当該複数の端末装置と通信可能に接続されたサーバ装置とから構成される通信システムにおける前記サーバ装置であって、前記複数の端末装置のうちの1つから送信された、複数の撮影対象を含む第1の画像と、当該第1の画像を構成する1または複数の画素からなる複数の領域のそれぞれに対応する実行き情報を受信する受信手段と、 20

この受信手段で受信した前記第1の画像と前記実行き情報とを基に、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を抽出する抽出手段と、

この抽出手段で取得した前記複数の撮影対象の3次元的な形状と3次元的な位置関係とのうちの少なくとも1つを基に、前記第1の画像に広告画像を合成することにより第2の画像を生成する生成手段と、

前記第2の画像を前記複数の端末装置のうちの他の端末装置へ送信する送信手段と、

を具備したことを特徴とするサーバ装置。 30

【請求項19】

少なくとも互いに通信可能な複数の端末装置と当該複数の端末装置と通信可能に接続されたサーバ装置とから構成される通信システムにおける前記サーバ装置であって、

前記複数の端末装置のうちの1つから送信された、少なくとも複数の撮影対象を含む第1の画像と、当該第1の画像を構成する1または複数の画素からなる複数の領域のそれぞれに対応する実行き情報から抽出された、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を含む特徴情報を受信する受信手段と、

この受信手段で受信した前記特徴情報に基づき、前記第1の画像に広告画像を合成することにより第2の画像を生成する生成手段と、

前記第2の画像を前記複数の端末装置のうちの他の端末装置へ送信する送信手段と、 40

を具備したことを特徴とするサーバ装置。

【請求項20】

複数の撮影対象を含む第1の画像を取得するとともに、1または複数の画素からなる複数の領域から構成される前記第1の画像の前記複数の領域のそれぞれに対応する実行き情報を取得し、

前記第1の画像と前記実行き情報とを基に、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を特徴情報として抽出して、この特徴情報を基に前記第1の画像を加工することを特徴とする画像加工方法。

【発明の詳細な説明】

【0001】

## 【発明の属する技術分野】

本発明は、例えば携帯電話などの端末装置において、撮影した画像に所望の画像を合成するなどして画像を加工する画像加工方法に関し、特に、この画像加工方法を用いて生成された画像を、携帯電話などの複数の端末装置間で送受信することでコミュニケーションを行う端末装置および通信システムに関する。

## 【0002】

## 【従来の技術】

近年、動画像を撮影することができるカメラ装置（ビデオカメラなど）は一般的に広く普及するようになった。これにより、自分の子供の成長過程や運動会などの行事といった身近な対象物を気楽に撮影・閲覧することが可能となった。また、最近では、携帯電話やPDAなどにもカメラが具備されていることが多く、街角などで気軽に動画像を撮影し、そのまま電子メールなどをもちいて、相手に動画像を添付したメールを送ることができるようになった。これらにより、ビデオ画像をコミュニケーションの手段として利用する活用方法が一般的になりつつある（例えば、非特許文献1、非特許文献2参照）。また、ネットワークのブロードバンド化も急速に進み、テレビ電話などといった、リアルタイムの映像コミュニケーションも既に実現されている（例えば、非特許文献3参照）。

## 【0003】

以上のように、一般の人々が気楽に動画像をコミュニケーションとして用いる文化が浸透しつつある。しかし、従来のビデオカメラでは、単純にカメラに写ったものを素材として用いることがほとんどで、それに対して、何らかの特殊効果を施すということはあまり行われていなかった。これは、人間の目とは異なり、従来のカメラが、対象の色情報のみを2次元的に撮影するものであったからである。そのため、人間の目では、カメラに写っている対象の区別や、その前後関係、立体形状などが把握できているのに対し、カメラに写ったものは、色情報でしか区別をつけることができない。

## 【0004】

このため、カメラで撮影した画像から、例えば、人間の部分のみを取り出して背景と区別したり、複数の人間の前後関係、位置関係を得たり、物の表面の形状変化を判別したり、といった処理を行うことは非常に困難であった。色情報のみを用いて、仮想的にそのような情報を得ようという試みもある。例えば、画像中からオブジェクトを背景から分離し、これを用いて、オブジェクトのみを用いて画像コミュニケーションを行う技術が開示されている（例えば、特許文献1参照）。ここでは、顔などのパーツを分離し、別に用意した背景に、対話をしている複数人の顔パーツを表示することで、サイバー空間でコミュニケーションを行っているような効果を与えている。しかし、本来、立体的な情報とは無関係の色情報を用いているため、厳密に対象の区別を行うことは困難、実際のオブジェクトの前後関係を把握しているのではない、など多くの制約が存在し、安定的に行うことは難しい。

## 【0005】

そこで、現状のビデオ画像を用いたコミュニケーションでは、撮影したビデオ画像に、このような対象の区別や前後関係などの位置関係、立体形状といった情報を用いて画像合成などの特殊効果をかけることは考えられていなかった。現在行われているのは、ビデオ画像の上に、文字を重ね書きしたり、飾りフレームを重ねたり、といった特殊効果がほとんどである（例えば、非特許文献4参照）。

## 【0006】

一方、上記個人間の映像コミュニケーションを拡張して、その映像中に広告を提示することが可能となれば、様々なビジネス展開を考えることができる。しかし、上述したように、従来の画像に広告の付加を考えた場合、どのように広告を重畳すればよいか、という問題がある。単に従来の画像の上にそのまま広告を重畳したのでは、邪魔になることが多い。例えば、人間のバストアップ画像を用いてテレビ電話によるコミュニケーションをしている最中に、自分や相手の顔の上に広告が重畳表示されると、邪魔になるだけでなく、ユーザがその広告に対して悪印象を持ってしまい、広告の意味がない。そこで、従来、画像

10

20

30

40

50

中に広告を付加するという試みは行われていなかった。

【0007】

【特許文献1】

特開2001-188910公報（段落番号「0158」乃至「0170」、第24図）

【0008】

【非特許文献1】

J-PHONE社のWWWページ

http://www.j-phone.com/movie-skamail/

【0009】

【非特許文献2】

NTTドコモ社のWWWページ

http://www.nttdocomo.co.jp/P\_S/imode/ishot/index.html

【0010】

【非特許文献3】

NTTドコモ社FOMAのWWWページ

http://foma.nttdocomo.co.jp/

【0011】

【非特許文献4】

ATLUSのWWWページ

http://www.atlus.co.jp/am/Printclub/

【0012】

【発明が解決しようとする課題】

このように、従来は、撮影した画像（動画像、静止画像を含む）に他の所望の画像を合成するなどして画像を加工する際、当該撮影した画像中の複数の撮影対象のそれぞれの3次元の形状や位置関係などを無視して、画像上の2次元平面上に単純に画像を合成するなどの加工を行うことしかできないという問題点があった。

【0013】

しかし、当該撮影した画像中の複数の撮影対象を区別して、それらの3次元の形状や位置関係などの情報を用いることができれば、これら画像中の撮影対象の3次元の情報を活用して、当該撮影した動画像中に映っている人間の周りにコンピュータグラフィックス（CG）のキャラクターが飛び回っていたり、背景部分のみを変化させたり、といったように、より変化に富んだ特殊効果を施すことが可能となる。

【0014】

また、背景と人物の部分の判別が可能となれば、広告画像を合成する際には、人間の部分に広告がかからないように背景部分に広告画像を合成するといったことも可能となる。さらに、撮影した動画像中に映っている人間の周りに広告を持ったコンピュータグラフィックスのキャラクターが飛び回るなど、対象の立体形状をうまく生かした新たな広告の提示方法も実現可能となる。

【0015】

そこで、本発明は、上記問題点に鑑み、従来技術では不可能であった、撮影した画像（動画像、静止画像を含む）中の複数の撮影対象のそれぞれの3次元の形状や3次元の位置関係などの3次元の特徴に合わせて、広告画像や所望の画像を合成するなどの加工が行える画像加工方法およびそれを用いた画像加工装置、通信端末装置およびサーバ装置を提供することを目的とする。

【0016】

【課題を解決するための手段】

（1）本発明は、互いに通信可能な複数の端末装置のうちの1つである端末装置であって、複数の撮影対象を含む第1の画像を取得する第1の取得手段と、前記第1の画像は1または複数の画素からなる複数の領域から構成され、この複数の領域のそれぞれに対応する

10

20

30

40

50

奥行き情報を取得する第2の取得手段と、前記第1の画像と前記奥行き情報とを基に、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係の特徴情報として抽出する抽出手段と、この抽出手段で抽出した特徴情報を基に、前記第1の画像を加工して第2の画像を生成する生成手段とを具備したことにより、撮影した第1の画像（動画像、静止画像を含む）中の複数の撮影対象のそれぞれの3次元的な形状や3次元的な位置関係などの3次元的な特徴に合わせて、所望の画像を合成するなどの加工が行える。

【0017】

例えば、前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該第1の画像中の奥行き方向に当該物体の位置を定めたときの、当該物体と前記複数の撮影対象の3次元的な位置関係と、前記複数の撮影対象の3次元的な形状とのうちの少なくとも1つを基に、当該物体の画像を当該第1の画像に合成する。

【0018】

また、前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、前記物体の動きが前記撮影対象の3次元的な形状に合うように制御して、当該物体の画像を当該第1の画像に合成する。

【0019】

また、前記抽出手段は、前記第1の画像中の前記撮影対象の存在する奥行き方向の位置と当該撮影対象の後の背景領域とを抽出し、前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該物体が当該撮影対象と前記背景領域の間に存在するように当該物体の画像を当該第1の画像に合成する。

【0020】

また、前記抽出手段は、前記第1の画像中の前記複数の撮影対象のうちの1つである第1の撮影対象の存在する奥行き方向の第1の位置と、当該第1の撮影対象の後にある、前記複数の撮影対象のうちの他の1つである第2の撮影対象の存在する奥行き方向の第2の位置とを抽出し、前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該物体が当該第1の位置と第2の位置の間に存在するように当該物体の画像を当該第1の画像に合成する。

【0021】

また、前記抽出手段は、前記第1の画像中の前記撮影対象の存在する3次元的な位置として第3の位置を抽出し、前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該物体の動きと当該撮影対象の動きとに基づき、当該物体が前記第3の位置に至ったとき、当該物体が当該撮影対象と衝突したと判断し、当該物体の動きや表現が衝突に対応するよう制御して当該物体の画像を当該第1の画像に合成し、また、前記物体が前記撮影対象と衝突したと判断したとき、衝突の効果表現を前記第1の画像に合成する。

【0022】

また、前記抽出手段は、前記第1の画像中の前記撮影対象の存在する3次元的な位置として第3の位置を抽出し、前記生成手段は、前記第1の画像に物体の画像を合成することにより前記第2の画像を生成するものであって、その際、当該撮影対象の前記第3の位置と、前記第1の画像中の当該物体の3次元的な位置である第4の位置とに基づき、当該物体の動きや表現を制御して当該物体の画像を当該第1の画像に合成する。

【0023】

また、前記抽出手段は、前記第1の画像から前記複数の撮影対象のそれぞれの画像と、それらの後の背景領域の画像とを抽出し、前記生成手段は、前記背景領域の画像と前記複数の撮影対象のそれぞれの画像のうちの少なくとも1つを加工することにより前記第2の画像を生成する。

【0024】

(2) 前記第2の取得手段は、前記第1の画像に対応し、各画素値に実行き情報を含む第3の画像を取得する手段と、前記第3の画像から前記第1の画像の各領域に対応する実行き情報を抽出する手段と、を具備し、前記抽出手段は、前記第1の画像と当該第1の画像の各領域に対応する前記実行き情報とから、少なくとも前記撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を抽出する。

【0025】

(3) 前記第2の取得手段は、前記撮影対象に向けて発光する発光手段と、この発光手段を発光させて前記撮影対象に照射した光の当該撮影対象からの反射光を含む第1の光量を受光する第1の受光手段と、前記発光手段が発光していないときに、前記反射光を含まない第2の光量を受光する第2の受光手段と、前記第1の光量から前記第2の光量を差し引いて、前記第1の光量から前記反射光の成分を抽出することにより、各画素値に実行き情報を含む第3の画像を生成する手段と、前記第3の画像から前記第1の画像の各領域に対応する実行き情報を算出する手段とを具備し、前記抽出手段は、前記第1の画像と当該第1の画像の各領域に対応する前記実行き情報とから、少なくとも前記撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を抽出する。

【0026】

(4) 前記生成手段は、前記第1の画像に衝撃などの効果表現や物体を表現した付加画像を合成することにより前記第2の画像を生成するものであって、この生成手段で前記第1の画像に合成するための複数種類の付加画像を記憶する記憶手段と、この記憶手段に記憶された複数の付加画像の中から前記第1の画像に合成する付加画像を選択する手段とをさらに具備したことにより、ユーザは、所望の付加画像を自由に選択して、自分の好みにあった特殊効果を上記第1の画像に付加することができる。

【0027】

(5) 前記生成手段は、前記第1の画像に衝撃などの効果表現や物体を表現した付加画像を合成することにより前記第2の画像を生成するものであって、この生成手段で前記第1の画像に合成するための複数種類の付加画像を記憶する記憶手段と、前記第1の画像から、当該第1の画像に写っている物体、色的な雰囲気、画像の構図のうちの少なくとも1つをシーンの特徴として抽出する手段と、前記シーンの特徴を基に、前記記憶手段に記憶された複数種類の付加画像の中から前記第1の画像に合成する付加画像を選択する手段とをさらに具備したことにより、上記第1の画像中の物体や色的な雰囲気、画像の構図などに適した特殊効果を付加することができる。

【0028】

(6) 本発明は、互いに通信可能な複数の端末装置と当該複数の端末装置と通信可能に接続されたサーバ装置とから構成される通信システムにおける前記サーバ装置であって、前記複数の端末装置のうちの1つから送信された、複数の撮影対象を含む第1の画像と、当該第1の画像を構成する1または複数の画素からなる複数の領域のそれぞれに対応する実行き情報を受信する受信手段と、この受信手段で受信した前記第1の画像と前記実行き情報とを基に、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を抽出する抽出手段と、この抽出手段で取得した前記複数の撮影対象の3次元的な形状と3次元的な位置関係との中の少なくとも1つを基に、前記第1の画像に広告画像を合成することにより第2の画像を生成する生成手段と、前記第2の画像を前記複数の端末装置のうちの他の端末装置へ送信する送信手段と、を具備したことにより、端末装置間で送受信される画像(上記第1の画像)に、当該第1の画像(動画像、静止画像を含む)中の複数の撮影対象のそれぞれの3次元的な形状や3次元的な位置関係などの3次元的な特徴に合わせて(当該第1の画像中の撮影対象のじゃまにならないように)広告画像を合成することができる。従って効率よく広告画像を合成することができ、広告効果の向上が図れる。

【0029】

また、上記サーバ装置は、前記複数の端末装置のうちの1つから送信された、少なくとも複数の撮影対象を含む第1の画像と、当該第1の画像を構成する1または複数の画素から

10

20

30

40

50

なる複数の領域のそれぞれに対応する奥行き情報から抽出された、少なくとも前記複数の撮影対象の3次元的な形状と前記複数の撮影対象の3次元的な位置関係を含む特徴情報を受信する受信手段と、この受信手段で受信した前記特徴情報に基づき、前記第1の画像に広告画像を合成することにより第2の画像を生成する生成手段と、前記第2の画像を前記複数の端末装置のうちの他の端末装置へ送信する送信手段とから構成されていてもよい。

【0030】

【発明の実施の形態】

以下、本発明の実施の形態について図面を参照して説明する。

【0031】

(第1の実施形態)

まず、本発明の第1の実施形態について説明する。

【0032】

<全体の構成>

図1は、第1の実施形態に係る、互いに通信可能な複数の端末装置のうちの1つである、主要部の構成例を示したものである。

【0033】

図1に示した画像コミュニケーション装置は、撮影対象の画像と当該画像の奥行き情報を取得する画像取得部1と、画像取得部1で取得した画像の奥行き情報をもとに当該画像中の撮影対象などの3次元的な特徴を抽出する特徴抽出部2と、特徴抽出部2で抽出された特徴をもとに、画像取得部1で取得した画像に3次元的な要素を考慮して、画像合成などの特殊効果を施す加工部3と、加工部3にて加工した結果得られた画像や、通信部5から受信した画像を提示する画像提示部4と、加工部3にて加工された画像を送受信するための通信部5とから構成される。

【0034】

図1に示したような構成を有する複数の端末装置間では、それぞれにおいて取得した画像に対し加工部3で加工を施して得られた画像を他の端末装置へ送信したり受信したりすること、各端末装置のユーザ間でコミュニケーションを行う。

【0035】

<画像取得部>

まず、画像取得部1について説明する。

【0036】

画像取得部1は、撮影対象の画像として例えばカラー画像を取得するとともに、当該画像の奥行き情報を取得して、撮影対象を、その3次元形状と画像取得部1から当該撮影対象までの距離を反映した奥行き情報を含むカラー画像（ここでは、奥行きカラー画像と呼ぶ）として取得するものである。

【0037】

なお、ここでは、撮影対象の画像としてカラー画像を取得する場合を例にとり説明するが、この場合に限らず、階調画像（いわゆる白黒画像）であってもよい。この場合には、画像取得部1では、撮影対象を、その3次元形状と画像取得部1から当該撮影対象までの距離を反映した奥行き白黒画像として取得する。

【0038】

さて、通常、カラー画像は、画素とよばれる単位で構成されており、画素には、図6(b)に示すように、例えばRGBの色の情報が格納されている。この画素を縦横の2次元の方向にアレイ状に並べることで実現されている。例えば、VGA(Video Graphics Array)サイズのカラー画像は、X軸(横)方向640画素、Y軸(縦)方向480画素の2次元アレイで表現されており、各画素内に、その位置における色の情報が例えばRGB形式で格納されている。なお、図6(b)は、X軸方向5画素、Y軸方向8画素のカラー画像において、各画素に格納されている色の情報(例えば、(R、G、B)=(r1、g1、b1)など)の一例を示したものである。

【0039】

10

20

30

40

50



奥行きカラー画像では、各画素に、例えば上記のようにRGB形式の色の情報に加え、その位置における奥行き情報（例えば、画像取得部1から撮影対象までの距離情報） $d$ が対応付けられている。ここでの対応付けとしては、例えば、奥行きカラー画像の各画素に、例えば、上記（R、G、B）にさらに、当該画素に対応する奥行き情報 $d$ を追加して、（R、G、B、 $d$ ）という形式で記憶するものとする。

【0040】

なお、カラー画像の代わりに、白黒画像を用いる場合には、当該白黒画像の画素値は、階調値となる。従って、この場合、奥行き白黒画像の各画素には、上記階調値に、さらに、当該画素に対応する奥行き情報 $d$ が記憶されている。

【0041】

また、ここでは、奥行きカラー画像の各画素が奥行き情報 $d$ を保持するとして説明するが、この場合に限らず、予め複数の画素からなる領域を単位領域として定めて、この単位領域毎に、奥行き情報 $d$ を対応付けるあるいは保持するようにしてもよい。

【0042】

なお、奥行きカラー画像の各画素にRGBなどの色情報に奥行き情報を追加したとしたとしても、奥行きカラー画像を表示する場合、この奥行き情報を用いなければ、奥行きカラー画像は通常のカラー画像と全く同様に表示される。各画素に対応する奥行き情報を用いることでカラー画像を3次元的に表示することもできる。

【0043】

図2に、通常の白黒画像と奥行き白黒画像の比較を示す。図2(a)が通常の白黒画像である。2次元的に各画素に階調値が納められている画像である。図2(b)に奥行き白黒画像の一例を示す。これは、奥行き情報を用いて3次元的に表現したものである。（3次元的な表現が分かりやすいように、正面下からの視点から見たとした。）このように、奥行き白黒画像および奥行きカラー画像は、従来の白黒画像やカラー画像とは異なり、奥行き情報として、画像取得部1から撮影対象までの距離情報 $d$ が含まれていることが特徴である。

【0044】

なお、ここでは、通常のカラー画像、白黒画像の各画素に奥行き情報に対応付けたものを、それぞれ、奥行き情報を含むカラー画像、奥行き情報を含む白黒画像とも呼ぶが、奥行きカラー画像、奥行き白黒画像は、それぞれ、奥行き情報を含むカラー画像、奥行き情報を含む白黒画像の一例という位置づけにある。

【0045】

それでは、画像取得部1の構成について図3を用いて詳細に説明する。図3は画像取得部1の機能ブロック図であり、大きく分けて、撮影対象の色情報が含まれる画像情報をリアルタイムで取得するための画像情報取得部101と、撮影対象の奥行き情報をリアルタイムで取得するための奥行き情報取得部102と、撮像動作制御部104と、奥行きカラー画像生成部113と、出力部114とから構成されている。

【0046】

画像情報取得部101は、撮影対象（例えば、物体など）を自然光（照明光を含む）で撮像することにより背景を含む物体のカラー画像である自然光画像を取得する自然光画像撮像部103と、自然光画像撮像部103で撮像された自然光画像が格納される自然光画像記憶部105と、自然光画像記憶部105に格納された自然光画像を読み出して、必要に応じて明度やコントラストなどの調整をする自然光画像調整部106と、自然光画像調整部106で調整された自然光画像を加工部3での処理に適したデータ形式で出力する加工部3へ出力する画像情報出力部107とから構成され、自然光画像調整部106で調整された自然光画像は、さらに、奥行きカラー画像生成部113へ出力される。

【0047】

自然光画像撮像部103は、例えば、CCDやCMOS撮像素子であり、背景を含む撮影対象（例えば物体など）のカラー平面画像をリアルタイムで取得する。これにより、物体の属性のうち色を取得できる。また、上記カラー画像を連続して撮像するので、連続する

10

20

30

40

50

フレームの変化から物体の動き情報を把握することができる。また、物体の画像を動画にすることができる。

【0048】

撮像動作制御部104は、自然光画像撮像部103と反射光画像撮像部108の動作を制御するための制御信号を発生する。

【0049】

一方、奥行き情報取得部102は、撮像動作制御部104からの制御信号に従って、物体やその周囲にある被写体などの撮影対象の反射光画像を撮像するための反射光画像撮像部108と、反射光画像撮像部108で撮像された反射光画像が格納される反射光画像記憶部109と、反射光画像記憶部109に格納された反射光画像を読み出して各種補正をする反射光画像補正部110と、反射光画像を補正するためのパラメータが記憶されているパラメータ記憶部111と、反射光画像補正部110で補正された反射光画像を解析して物体の奥行き情報を演算する奥行き情報演算部112とから構成され、奥行き情報演算部112で演算された奥行き情報は、奥行きカラー画像生成部113へ出力される。

【0050】

ここで、反射光画像撮像部108の詳細を図4を参照して説明する。図4は、反射光画像撮像部108の機能ブロック図である。反射光画像撮像部108は発光部114、受光光学系115、反射光抽出部116およびタイミング制御部117から構成される。

【0051】

発光部114はタイミング制御部117によって生成されるタイミング信号に従って時間的に強度変動する光を発生する。この光は発光部114の前方にある、例えば図4ではユーザの頭部（物体の一例）に照射される。ここで発生する光としては近赤外光であるが、これに限らず可視光など他の波長領域の光も利用することができる。

【0052】

例えばユーザの頭部からの反射光はレンズなどで構成される受光光学系115により集光されて反射光抽出部116の受光面上に結像される。受光光学系115には近赤外光を通過するフィルターが設けられている。このフィルターにより反射光のうち近赤外光以外の可視光や遠赤外光のような外光がカットされる。

【0053】

反射光抽出部116は、上記結像を形成する反射光の空間的な強度分布を抽出する。この強度分布は反射光による画像として捉えることができる。これが上記反射光画像（シルエット像）である。この機能を達成するため、反射光抽出部116は、第1の受光部119、第2の受光部120および差分演算部121から成る。第1の受光部119と第2の受光部120は、異なるタイミングで受光を行う。そして、第1の受光部119が受光しているときに発光部114が発光し、第2の受光部120が受光しているときには発光部114は発光しないように、タイミング制御部117がこれらの動作タイミングを制御する。これにより第1の受光部119は発光部114からの光の物体による反射光とそれ以外の自然光（つまり太陽光、照明光などの外光）を受光する。一方、第2の受光部120は自然光のみを受光する。両者が受光するタイミングは異なっているが近いので、この間における自然光の変動は無視できる。

【0054】

従って、第1の受光部119で受光した像と第2の受光部120で受光した像の差分をとれば、発光部114からの光のうち物体による反射光の成分だけが抽出され、反射光画像を生成することができる。差分演算部121が第1の受光部119と第2の受光部120で受光した像の差を計算して出力する。

【0055】

物体からの反射光は、物体と反射光画像撮像部108の距離が長くなるに従い大幅に減少する。物体の表面が一様に光を散乱する場合、反射光画像1画素あたりの受光量は物体までの距離の二乗に反比例して小さくなる。すなわち、例えば、反射光画像中の座標（ $i$ 、 $j$ ）にある画素の画素値を $Q(i, j)$ とすると、

10

20

30

40

50

$$Q(i, j) = K / d^2 \quad (1)$$

と表すことができる。ここで、 $K$ は、例えば、 $d = 0.5 \text{ m}$ のときに、 $Q(i, j)$ の値が「255」になるように調整された係数である。式(1)を $d$ について解くことで、距離を求めることができる。

【0056】

このように、反射光画像の各画素には、奥行き情報として、反射光画像撮像部108から撮影対象までの距離に換算することのできる反射光の強度値、すなわち、奥行き情報が含まれているのである。

【0057】

物体からの反射光の強度値は、反射光画像撮像部108から物体までの距離に換算できるので、物体の立体形状を把握することができる。また、背景からの反射光はほぼ無視できくくらいに小さい。よって、背景がカットされた、物体およびその周囲にある被写体の反射光画像を得ることができる。

【0058】

反射光画像の一例を図6(a)に示す。図6(a)には、簡単のため、 $256 \times 256$ 画素の反射光画像の一部である $5 \times 8$ 画素の反射光画像の場合について示している。各画素の画素値は、反射光の強度値である。

【0059】

人間の手を撮影対象として、反射光画像撮像部108で撮影された反射光画像に含まれる上記奥行き情報から得られる、当該反射光画像中の手の3次元的なイメージを図5に示す

【0060】

反射光画像は、例えば、以下の(1)から(3)の構成を有する装置により得ることができる。(1)時間的に一定あるいは時間的に変化するパルス信号や変調信号を発生させるためのタイミング信号生成手段。(2)このタイミング信号生成手段によって生成された信号に基づいて、強度変化する光を発するための発光手段。(3)この発光手段から発された光の物体による反射光をタイミング信号生成手段からの信号と同期して外光(自然光)から分離して検出する手段を配列して構成し、光の物体による反射光画像を検出する反射光抽出手段。

【0061】

反射光画像撮像部108の上記構成要素のさらに詳細な説明は、本件と同一出願人が出願した特開平10-177449号公報に記載されている。

【0062】

なお、反射光画像撮像部108は、物体を含む被写体の反射光画像を連続して撮像するので、連続するフレームの変化から物体の動き情報を把握することができる。これにより、物体の画像を動画にすることができる。

【0063】

後述するように、自然光画像(背景を含む物体のカラーの平面画像)と、物体の反射光画像を演算して得られる奥行き情報とを組み合わせて、奥行きカラー画像が生成される。このため、反射光画像の画素と自然光画像の画素とを対応させる必要がある。反射光画像撮像部108と自然光画像撮像部103は近接して配置するのが好ましい。この観点からこれらの自然光画像撮像部103、反射光画像撮像部108、さらに、必要に応じて撮像動作制御部104は1チップ化されていることが好ましい。

【0064】

また、自然光画像と反射光画像との間では、例えば画素間あるいは複数の画素からなる領域間の対応付けがなされていることが必要である。例えば、画像取得部101で取得される自然光画像(カラー画像)が図6(b)に示したような画像であり、奥行き情報取得部102で取得される反射光画像が図6(a)に示したような画像であるとする。この場合、カラー画像の各画素と反射光画像の各画素との間には、1対1に対応付けられている。

図6(a)と図6(b)の各画像中の各画素をその座標( $i, j$ )を用いて表すとき(こ

10

20

30

40

50

ここで、 $i = 1 \sim 5$ 、 $j = 1 \sim 8$ ）、図6(a)中の反射光画像の画素P2( $i$ 、 $j$ )と、図6(b)の自然光画像中の画素P1( $i$ 、 $j$ )とは互いに対応する画素となる。例えば、図6(a)中の反射光画像の画素P2(5、8)、図6(b)の自然光画像中の画素P1(5、8)とは互いに対応する画素となる。従って、奥行きカラー画像中の画素P3( $i$ 、 $j$ )の画素値は、自然光画像中の画素P1( $i$ 、 $j$ )の色情報と、反射光画像の画素P2( $i$ 、 $j$ )から(奥行き情報演算部112にて)算出された距離(奥行き情報) $d$ とが含まれている。

#### 【0065】

ここでは、自然光画像の画素と反射光画像の画素とが対応付けられている場合を示したが、この場合に限らず、自然光画像と反射光画像のサイズや解像度などの違いから、例えば、自然光画像を複数の画素からなる複数の単位領域に分割し、この複数の単位領域のそれぞれと反射光画像の各画素との間で対応付けを行うようにしてもよいし、また逆に、反射光画像を複数の画素からなる複数の単位領域に分割し、この複数の単位領域のそれぞれと自然光画像の各画素との間で対応付けを行うようにしてもよい。さらに、反射光画像と自然光画像のそれぞれを複数の画素からなる複数の単位領域に分割し、反射光画像と自然光画像のそれぞれの単位領域間で対応付けを行うようにしてもよい。

#### 【0066】

また、撮像動作制御部104は、例えば、ユーザからの撮像指示が入力されたときに、自然光画像撮像部103と反射光画像撮像部108とでほぼ同時に自然光画像と反射光画像とを取得するよう、自然光画像撮像部103と反射光画像撮像部108とに制御信号を出

#### 【0067】

次に、図3に示す反射光画像補正部110での補正を説明する。物体の色や反射特性などを考慮して補正をする。詳しく説明すると、物体からの反射光の強度は、物体と反射光画像撮像部108との距離以外の要因にも左右される。このため、反射光画像から単純に距離を求めても、距離(つまり立体形状)が正確でないことがある。例えば、物体表面の色が黒い場合、反射光の強度は低下する。また、物体の表面が鏡面反射成分を多く含む場合、物体の表面の法線が光源方向に近くなる部分で強い反射光が発生する。

#### 【0068】

よって、物体の反射光強度から距離情報を求める前に、反射光画像補正部110において、パラメータ記憶部111に予め格納されている物体の表面の色や反射特性などに関するパラメータを参照したり、自然光画像記憶部105に格納されている自然光画像を参照して、当該自然光画像中の反射光画像の画素や単位領域に対応する領域(画素や単位領域)の画素値を基に、物体の反射光画像を補正する。例えば、反射光画像中のある画素に対応する自然光画像の画素の色情報が「黒」であれば、反射光画像の当該画素値を黒い撮影対象からの反射光を受光した場合のパラメータを用いて補正する。

#### 【0069】

これにより、反射光画像撮像部108で得られた反射光画像の各画素値が補正されるので、後段の奥行き情報演算部112で算出される奥行き情報の精度を高めることができる。

#### 【0070】

奥行き情報演算部112は、反射光画像補正部110で補正された反射光画像の各画素について、その画素値としての受光量(反射光強度)を $Q$ とすると、当該画素に対応する距離 $d$ を、例えば次式(2)から求める。

#### 【0071】

$$d = (K/Q)^{1/2} \quad (2)$$

すなわち、前述したように、物体からの反射光の強度値は、反射光画像撮像部108から物体までの距離 $d$ に換算できるので、この距離 $d$ を奥行き情報として求めるのである。

#### 【0072】

なお、ここでは、反射光画像中の各画素について、奥行き情報としての距離 $d$ を求める場合を示したが、この場合に限らず、例えば、上記単位領域毎に、その代表画素あるいは当

10

20

30

40

50

該単位領域中の画素値の平均値などから、距離 $d$ を算出するようにしてもよい。

【0073】

さて、奥行きカラー画像生成部113では、自然光画像調整部106から出力された自然光画像と、奥行き情報演算部112で算出された反射光画像中の各画素（あるいは各単位領域）の奥行き情報とから、奥行きカラー画像を生成する。

【0074】

すなわち、例えば、自然光画像である、例えば図6（b）に示したようなカラー画像の各画素の画素値に、当該画素に対応する反射光画像の画素から算出された奥行き情報を追加して、奥行きカラー画像の各画素値を生成する。例えば、図8（a）中の反射光画像の画素P2（5、8）と、図6（b）の自然光画像中の画素P1（5、8）とは互に対応する画素となる。従って、自然光画像中の画素P1（5、8）の画素値に、反射光画像の画素P2（5、8）から（奥行き情報演算部112にて）算出された距離（奥行き情報） $d$ とを追加して、奥行きカラー画像の画素P3（5、8）の画素値を生成する。

【0075】

なお、自然光画像を複数の画素からなる複数の単位領域に分割し、この複数の単位領域のそれぞれと反射光画像の各画素との間に対応付けを行っている場合には、自然光画像の各単位領域に、当該単位領域に対応する反射光画像の画素から算出された奥行き情報 $d$ を対応付けて（例えば、当該単位領域中の各画素の画素値に当該単位領域に対応する反射光画像の画素から算出された奥行き情報 $d$ を追加する）、奥行きカラー画像の各画素値を生成する。

【0076】

また、逆に、反射光画像を複数の画素からなる複数の単位領域に分割し、この複数の単位領域のそれぞれと自然光画像の各画素との間に対応付けを行っている場合には、自然光画像の各画素の画素値に、当該画素に対応する反射光画像の単位領域から算出された奥行き情報を追加して、奥行きカラー画像の各画素値を生成する。

【0077】

さらに、反射光画像と自然光画像のそれぞれを複数の画素からなる複数の単位領域に分割し、反射光画像と自然光画像のそれぞれの単位領域間に対応付けを行っている場合には、自然光画像の各単位領域に、当該単位領域に対応する反射光画像の単位領域から算出された奥行き情報 $d$ を対応付けて（例えば、当該単位領域中の各画素の画素値に当該単位領域に対応する反射光画像の単位領域から算出された奥行き情報 $d$ を追加する）、奥行きカラー画像の各画素値を生成する。

【0078】

このようにして生成された奥行きカラー画像は、出力部114に送られ、ここで、出力先のプロトコルやデータ形式に合わせるために変換等を行って、特徴抽出部2へ出力される。

【0079】

なお、以上説明した画像取得部1の構成は、あくまでも一例であり、これに限定されるものではない。特に、奥行き情報を取得する際には、上記のように、反射光画像を必ずしも用いる必要はない。すなわち、複数の視野から撮影した画像の視差情報を用いることで奥行き情報を計算するという、ステレオマッチングの手法を用いて奥行き情報を取得するという構成であってもよいし、縞状のレーザー光を照射し、その形のやがみを用いて上記奥行き情報を計測するというレーザーレンジファインダと呼ばれる方式を用いてもよい。また、これら以外の方法を用いて奥行き情報を取得して、上記のような奥行きカラー画像を取得することができるものを使用することでもできる。

【0080】

また、上記画像取得部1は、奥行きカラー画像生成部113で各画素に奥行き情報を含む画像（奥行きカラー画像）を生成するようになっているが、この場合に限らず、自然光画像の各画素に、当該画素に対応する奥行き情報を対応付けて、その対応関係を保持あるいは記憶するだけでもよい。

【0081】

<特徴抽出部>

次に、特徴抽出部2について説明する。ここでは、画像取得部1で求めた画素値に奥行き情報を含む奥行きカラー画像を処理対象とする。

【0082】

特徴抽出部2は、画像取得部1で取得した奥行きカラー画像に含まれる奥行き情報をもとに撮影対象の3次元的な特徴を抽出するためのものである。

【0083】

ここで、図7に示す内容の奥行きカラー画像を参照して、特徴抽出部2で解析する3次元的な特徴に関して具体的に説明する。なお、図7は、奥行きカラー画像に含まれる奥行き情報を用いずに表示した場合を示しているのも、通常のカラー画像（図7の場合、白黒画像）と同様である。図7は、シーン中央に、イスに座った人物の上半身が映っており、その人物は、右腕を挙げ、人差し指を立てている。また、人物の手前に缶ジュースが置かれている。

10

【0084】

このシーンにおけるそれぞれの物体の位置関係を模式的に簡潔に示したのが図8である。図8のように、画像取得部1から最も近い距離に缶が存在し（図12参照）、その向こうに、人物が存在している（図10参照）。さらに遠くに背景部分がある（図9参照）。また、人物内の奥行き情報を細かく見てみると、右腕部分が胴体よりも画像取得部1から近い部分にある（図11参照）。つまり、奥行きの違いによって、シーンを図9～図12に示すように、幾つかの領域に分割することができるわけである。

20

【0085】

このように、画像取得部1で取得した奥行きカラー画像に含まれる奥行き情報を解析することで、シーン内にどのような撮影対象（例えば物体）が存在するか（どのような領域に分割できるか）、シーン内の奥行き方向の凹凸関係はどうなっているか、シーンにおける各物体（1つの物体を構成する部分もそれぞれ1つの物体としてみなす）の3次元的な位置関係、それぞれの物体の立体形状、などといった特徴を得ることができる。

【0086】

それでは、特徴解析手法に関して説明する。

【0087】

最も簡単な解析として、シーン内部に撮像されている物体が何かを判別せずに、その3次元形状のみを得るというものがある。例えば、図7のシーンの場合、画面中央下部（実際には、缶の部分）が最も近い奥行き値を持ち、画面左側下部（実際には人物の右腕部分）が次々に近い奥行き値を持ち、画面中央の大部分（実際には、人物部分）が次々に存在し、その後の部分（実際には背景部分）は、非常に遠くに存在する、などといったように、奥行き方向に関してシーンの全体的な凹凸関係を抽出する。

30

【0088】

また、特徴抽出部2では、必要ならば、さらに複雑な解析を行うことも可能である。

【0089】

図13は、図7に示した奥行きカラー画像に含まれる奥行き情報を模式的に示したものである。図13は、図7に示したシーンを上方から見て、奥行き情報の得られている部分のみを実線で示したものである。このように、奥行きカラー画像では、撮影方向から見える部分の奥行き情報のみが得られる。（撮影方向から見えない部分、例えば、この例では、缶や人物の後ろ側、人物の右腕の後ろに隠れた部分などのデータは無い。）このような奥行き情報をもとに撮影対象の3次元的な位置関係を得る。

40

【0090】

最も簡単な方法は、画像取得部1からの奥行き情報 $d$ によって単純に前景と背景を区別することである。ある値 $TH$ を定め、奥行き情報が値 $TH$ よりも近い部分を前景、遠い部分を背景とみなす（図14参照）。このようにすることで、シーンに映っているものと背景部分を区別することが可能である。これにより、撮影しているシーンに、何か物体（

50

つまり前景にあたる部分)が存在しているか、を判別することができる。

【0091】

また、この 値を複数用意し、例えば奥行きカラー画像に写っているシーンを第1の 値 $\times 1$ と第2の 値 $\times 2$ と第3の 値 $\times 3$ とで、距離 $\times 1$ より手前と、距離 $\times 1$ と $\times 2$ の間、距離 $\times 2$ と $\times 3$ の間、距離 $\times 3$ 以降というように、4つの領域に分割し、各領域について、物体が存在するか否か、その物体は何であるかを識別したりするようにしてもよい。

【0092】

また、ジャンプエッジと呼ばれる奥行き方向の不連続点(図15における点線で示された部分)を検出することで、図15に示すように、「背景部分」、「人物部分」、「人物の右腕部分」、「缶の部分」、などといったように、シーン内のさらに細かい位置関係を認識することも可能である。

【0093】

ジャンプエッジの検出方法は、様々な手法を用いることができるが、例えば、奥行きカラー画像から、奥行き部分のみを抽出した画像(反射光画像などの、奥行き情報のみが2次元のアレイドに並んでいるデータで、以降、奥行き画像と呼ぶ)に対して、エッジ検出のためのフィルタリング処理(代表的には、Sobelオペレータによる畳み込みフィルタリング処理など)をおこなうことで、得ることができる。

【0094】

さらに、パターンマッチングという、あらかじめ物体の特徴を登録しておき、画像内にその特徴と類似している部分を探す手法があるが、その手法などを用いることで、上記で得られたそれぞれの物体が何であるかを認識することができる。例えば、図7に示したシーン内において、「缶」「人物」、当該人物の「右手」などの撮影対象が認識でき、しかも、それらの位置関係が認識できれば、図7からは、「缶」があり、その「缶」の位置はどこであるか、「缶」の後には、「人物」がいて、「右手」を挙げている、などといったような複雑な認識も可能である。

【0095】

なお、上記のような、奥行きカラー画像から3次元的な特徴を抽出するための解析手法は、あくまでも一例であり、これに限定されるものではない。他の様々な解析・画像処理・画像認識の手法を組み合わせて実現することが可能である。

【0096】

特徴抽出部2では、奥行きカラー画像(自然光画像と、反射光画像の各素値から得られた奥行き情報)から、自然光画像中の3次元的な特徴を抽出する。自然光画像中の3次元的な特徴とは、例えば、自然光画像中に各撮影対象の3次元的な形状(表面上の凹凸状態も含む)、複数の撮影対象の位置関係(自然光画像中の平面方向の位置関係と、自然光画像の奥行き方向の位置関係(主に前後関係))などであり、さらに、これらから、各撮影対象に対応する奥行き方向の位置から(予め定められた 値に基づき)判別された撮影対象の存在する前景部分と、背景部分、さらに細かな領域分割が行えるときともに、パターンマッチングなどにより撮影対象が何であるかを認識することもできる。

【0097】

<加工部>

次に、加工部3について説明する。

【0098】

加工部3は、特徴抽出部2で抽出された3次元的な特徴をもとに、画像取得部1で取得した自然光画像であるカラー画像(あるいは奥行きカラー画像)に、当該カラー画像中の撮影対象の3次元的な形状や位置関係などの3次元的な特徴を考慮した特殊効果を施す(付加する)ためのものである。

【0099】

具体的には、カラー画像に、CG(コンピュータグラフィックス)で表現された画像(ここでは、付加画像ともいう)を合成することで、特殊効果を付加する。この際、特徴抽出部2で抽出された、カラー画像(シーン)中の凹凸といった奥行き情報、カラー画像中に

10

20

30

40

50

どのような物体が存在するか（どのような領域に分割できるか）、カラー画像中における各物体の位置関係、それぞれの物体の立体形状、などといった3次元的な特徴を活用することで、仮想物体とシーンの前後関係や衝突状態などを判別し、必要に応じて仮想物体を変形し、カラー画像に合成する。

#### 【0100】

ここで、図7に示したカラー画像（に写っているシーン）を例として、3D（3次元）CGのデータとして与えられる仮想物体「球」を合成する場合を考える。図16は、図7のカラー画像中の撮影対象である各物体の主に奥行き方向の位置関係を示したものであるが、上述したように、このようなシーンにおける3次元的な特徴が、特徴抽出部2から得られている。

10

#### 【0101】

いま、「球」が図16における奥行き位置Cのところを、画面右から左に動くという特殊効果を考える。この際、仮想物体（仮想オブジェクト）である「球」の置かれる3次元的な位置および、その形状情報は既知であるため、これと、シーンの3次元的な特徴（シーンの各位置における奥行き値）を比較することで、「球」と、シーン中の各物体の位置との前後関係を判別することができる。

#### 【0102】

これより、図17に示すように、球が背景（図16における奥行き位置D）の前を通るが、人物（奥行き位置B）の後ろを通るように、「球」を合成することが可能である。このように、特徴抽出部2で抽出された特徴を基に、カラー画像に3次元的に仮想物体を付加

20

#### 【0103】

同様に、カラー画像中の奥行き位置Aのところを、「球」が右から左へ移動するように、仮想物体「球」の画像を合成すると、「球」は、缶の後ろを通り人物の前を通る特殊効果となる。

#### 【0104】

さらに、これを推し進め、仮想物体とカラー画像中の物体との衝突判定も行うことができる。なお、図16からも明らかなように、カラー画像中の物体の位置は、画像平面上と奥行き方向とから3次元的に特定することができる。従って、カラー画像中の各物体のそれぞれについて特定される3次元的な位置を基により精密な特殊効果を付けることもできる。その1つが「衝突」の特殊効果である。

30

#### 【0105】

いま、図16における奥行き位置Bのところに、画面右側から仮想物体「球」が移動してくる特殊効果を考える。この際、特徴抽出部2で抽出された特徴によれば、人物のいる部分が奥行き位置Bであることが分かるため、「球」は、人物の位置まで来たときに、人物と衝突することが分かる。そこで、図18に示すように、「球」が、人物と衝突し、跳ね返る、といった特殊効果も付加することができる。この際、衝突したあとに跳ね返る方向をも、物体の3次元的な形状を見ることができ、計算することが可能である。さらに、衝突する際に、図18のように、カラー画像中の物体の衝突箇所（図18では人物の頬の部分）に「衝突」の効果表現の1つである「星」を表現する、といった特殊効果を付加すると、より効果的である。

40

#### 【0106】

また、カラー画像中の各物体のそれぞれについて特定される3次元的な位置を基に、次のような特殊効果も可能である。例えば、仮想物体が3次元的に動く特殊効果の場合について説明する。図19に、カラー画像手前から当該画像の奥行き方向側に、2つの仮想物体「球」が移動する特殊効果の例を示している。図19に示したように、特徴抽出部2で抽出された特徴によれば、缶が奥行き方向手前にあり、その後方に人物がいることが分かっている。この情報を「球」の動きに組み合わせ、図19に示したように、2つの「球」のうち、1つは、近くの缶に当たって跳ね返り、もう1つは、遠くの人物に当たって跳ね返る、といった特殊効果を付加することができる。

50



## 【0107】

また、図20に示すように、仮想物体「球」を人物の頭の周りを回るような特殊効果を付加したり、図21に示すように、「球」を指の周りを回るような特殊効果を付加したりすることも可能である。

## 【0108】

さらには、カラー画像中の物体の3次元的な形状、すなわち、凹凸という3次元的な特徴を用いることで、図22に示したように、人物の上から、ペンキなどの粘性のある液体などをたらす、といったような特殊効果をかけることもできる。図22のように、人物の鼻の部分は他の顔よりも出っ張っているなどの特徴があるため、鼻の部分を避けてペンキが流れていくといったことが再現できるのが特徴である。

10

## 【0109】

また、カラー画像中の物体の3次元的な位置関係の特徴も分かっているため、ペンキは、右腕の指の部分や、缶の部分に流れないように、ペンキという仮想物体の動きや形状を制御できる。

## 【0110】

さらに、缶の上や人間の肩の部分に仮想の「キャラクタ」を座らせたり、指の上の部分に仮想の「とんぼ」がとまったりといったように、各種の画像をカラー画像中の奥行き方向の所望の位置に合成することも可能である。

## 【0111】

また、図23に示すように、画面中央に仮想の「球」があって、それに対して、手を動かすことで、「球」が動く、というように、動画像のなかで、時間的に変化する手の動きなどの3次元的な特徴の情報を用いて、例えば、当該手の動きに合わせて「球」に仮想物に動きを与える（この場合、「球」を手で叩いてとばす）といった特殊効果も与えることができる。

20

## 【0112】

以上では、カラー画像に、仮想物体の画像を合成する特殊効果を例として説明したが、特殊効果は、これだけではない。例えば、背景部分を消して、別の背景に差し替えたり、CGで作成した背景に置き換えたりという画像合成による特殊効果も可能である。

## 【0113】

また、前景にある物体（図7の例では、人物と缶の部分）はそのままにして、背景部分のみを白黒やセピア色に変色させるなどの特殊効果も可能である。逆に、人物部分のみにモザイクをかけるなどということもできる。さらに、特殊効果は、カラー画像の全部または一部の変形を行うことで実施されてもよい。例えば、人物の顔の部分を風船のように膨らませたり、萎ませたり、といった特殊効果も可能である。

30

## 【0114】

以上説明したように、加工部3では、特徴抽出部2で抽出された3次元的な特徴をもとに、画像取得部1で自然光画像として取得したカラー画像に3次元的な要素を考慮した様々な特殊効果を付加する。

## 【0115】

特殊効果の一例としては、例えば、カラー画像に仮想物体を合成するときには、当該カラー画像中の奥行き方向の位置を定めることができるので、3次元空間内のその奥行き方向の位置に実際に当該仮想物体が存在するように、当該仮想物体の画像を当該カラー画像に合成する。その際、仮想物体の画像が動画の場合であっても、カラー画像が動画の場合であっても同様である。

40

## 【0116】

また、カラー画像中の物体の動きと位置や、仮想物体の動きと位置から、カラー画像中の物体と仮想物体の衝突を判定することができるので、この衝突に対応する仮想物体の動きを表現したり、衝突の発生したことを表す表示を行う（例えば、衝突を表す「星」のような仮想物体の画像を合成する）こともできる。

## 【0117】

50

また、カラー画像に仮想物体を合成するときには、当該カラー画像中の物体の3次元形状、すなわち、例えば凹凸に合わせて仮想物体の動きや形状を制御することができるので、3次元空間内に実際に当該仮想物体が存在するように、当該仮想物体の画像を当該カラー画像に合成する。その際、仮想物体の画像が動画の場合であっても、カラー画像が動画の場合であっても同様である。

【0118】

また、カラー画像中の撮影対象の物体や背景部分が3次元的な特徴として抽出されているので、例えば、撮影対象の物体や背景部分をそれぞれ別個に白黒やセピア色などに変色させたり、モザイクをかけたり、変形したりなどといったことも行える。

【0119】

なお、本実施形態で説明した特徴の使い方や特殊効果はあくまでも一例であり、これに限定されるものではない。

【0120】

また、ここでは、便宜上、「球」などの仮想物体を用いて説明したが、これに限定されるものではない。仮想物体としては、キャラクタ、乗り物、建物などの様々なものが考えられる。また、仮想物体といっても、実際の写真を用いて3DCGのデータにしたもの、ユーザにより入力された手書きのマークやキャラクタなどをCGデータにしたものも含む。また、本実施形態では、仮想物体が変形しないものとして説明したが、これに限定される物ではなく、仮想物体の種類に応じて、自由に変形しても構わない。また、仮想物体の画像が動画であってもよい。

【0121】

なお、ユーザが手書きのマークやキャラクタや乗り物、建物、その他様々なものの絵を入力するための入力部を図1に示した構成に新たに追加してもよい。そして、この入力部を通じて入力された手書きの絵を上記加工部3で処理可能なように、CGデータ化あるいは3DCGデータ化するための処理部も必要となる。

【0122】

以上説明したように、加工部3では、奥行きカラー画像（奥行き情報を含まないカラー画像であってもよい）に仮想物体の画像を合成する場合には、当該奥行きカラー画像中の奥行き方向に当該仮想物体の位置を定めたときの、当該仮想物体と撮影対象の3次元的な位置関係と、撮影対象の3次元的な形状とのうちの少なくとも1つを基に、当該仮想物体の画像を当該奥行きカラー画像に合成する。

【0123】

例えば、仮想物体の動きが撮影対象の3次元的な形状に合うように制御して、当該仮想物体の画像を当該奥行きカラー画像に合成する。

【0124】

また、特殊抽出部2で奥行きカラー画像中の撮影対象の存在する奥行き方向の位置と当該撮影対象の後の背景領域とを抽出し、加工部3では、仮想物体が撮影対象と背景領域の間に存在するように当該仮想物体の画像を当該奥行きカラー画像に合成する。

【0125】

また、特殊抽出部2で奥行きカラー画像中の複数の撮影対象のうちの1つである第1の撮影対象の存在する奥行き方向の第1の位置と、当該第1の撮影対象の後にある、上記複数の撮影対象のうちの他の1つである第2の撮影対象の存在する奥行き方向の第2の位置とを抽出し、加工部3では、仮想物体が当該第1の位置と第2の位置の間に存在するように当該仮想物体の画像を当該奥行きカラー画像に合成する。

【0126】

また、特徴抽出部2で、奥行きカラー画像中の撮影対象の存在する3次元的な位置として第3の位置を抽出し、加工部3では、仮想物体の動きと撮影対象の動きとに基づき、当該仮想物体が上記第3の位置に至ったとき、当該仮想物体が当該撮影対象と衝突したと判断し、当該物体の動きや表現が衝突に対応するよう制御して当該仮想物体の画像を当該奥行きカラー画像に合成する。なお、仮想物体が撮影対象と衝突したと判断したときには、す

10

20

30

40

50

りに衝突の効果表現（例えば、「星」など）を当該奥行きカラー画像に合成してもよい。

【0127】

また、特徴抽出部2で奥行きカラー画像中の撮影対象の存在する3次元的位置として第3の位置を抽出し、加工部3では、当該撮影対象の上記第3の位置と、上記奥行きカラー画像中の仮想物体の3次元的位置である第4の位置とに基づき、当該仮想物体の動きや表現を制御して当該仮想物体の画像を当該奥行きカラー画像に合成する。

【0128】

また、特徴抽出部2で、奥行きカラー画像から複数の撮影対象のそれぞれの画像と、それらの後の背景領域の画像とを抽出し、加工部3では、上記背景領域の画像と複数の撮影対象のそれぞれの画像のうちの少なくとも1つを加工する。例えば、撮影対象の物体や背景部分をそれぞれ別個に白黒やセピア色などに変色させたり、モザイクをかけたり、変形したりなどとする。

【0129】

<画像提示部>

次に、画像提示部4について説明する。

【0130】

画像提示部4は、加工部3にて特殊効果の施されたカラー画像（以下、特殊効果付き画像と呼ぶ）、通信部5にて受信した特殊効果付き画像をユーザに提示するためのものである。

【0131】

画像提示部4は、具体的には、ディスプレイ装置で構成され、加工部3にて生成された特殊効果付き画像をディスプレイ上に表示する。また、通信部5にて受信した特殊効果付き画像をディスプレイ上に表示する。

【0132】

さらに、図24に示したように、ディスプレイ装置の表示画面上の表示エリアを、加工部3にて生成された特殊効果付き画像を表示するエリアと、通信部5にて受信した特殊効果付き画像を表示するエリアとに分割し、双方を同時に提示することも可能である。図24では、エリアA1には、例えば、通信部5にて受信した特殊効果付き画像を表示され、エリアA2には、例えば、加工部3にて生成された特殊効果付き画像が表示される。

【0133】

また、画像提示部4は、加工部3にて生成された特殊効果付き画像を、奥行き情報を用いることで3次元モデル化して、3D（3次元）のシーンとして提示することも可能である。3Dのシーンにしてしまえば、視点の位置を変えて見たり、立体視をしたりすることも可能となる。

【0134】

<通信部>

最後に、通信部5について説明する。

【0135】

通信部5は、加工部3にて生成された特殊効果付き画像を、他の端末装置へ送信したり、他の（例えば、図1と同様な構成を有する）端末装置から送信されてきた上記同様の特殊効果付き画像を受信する。

【0136】

通信部5は、有線の通信手段による場合と、無線の通信手段による場合がある。まず、無線の通信手段による場合について説明する。

【0137】

この場合、例えば、PDC（Personal digital cellular）やCDMA（Code-Division Multiple Access）、PHSといった携帯電話に用いられているような無線通信方式を用いて、他の端末装置などの外部機器と通信する。これにより、外部機器への特殊効果付き画像の送信、外部機器からの特殊効果付き画像の受信が行われる。なお、通信手段は、携帯電話通信に限定されるもので

10

20

30

40

50

はなく、IEEE 802.11a/b/gなどに規定された無線LANや、Bluetooth（商標）、赤外線通信、RF通信、その他の無線通信方式を用いることが可能である。

【0138】

次に、通信部5が、有線の通信手段による場合について説明する。

【0139】

この場合、通信部5はUSB、IEEE 1394などのインタフェースを具備し、これらの方式で、接続された外部機器と通信を行う。そして、外部機器への特殊効果付き画像の送信、外部機器からの特殊効果付き画像の受信が行われる。例えば、USBによって、接続されたPCへ特殊効果付き画像を送信したりなどである。なお、通信手段は、これに限  
10  
定されるものではなく、シリアル通信、一般電話網、光ファイバ、その他の方式を用いることが可能である。また、外部機器の持っている通信手段を介して、さらに別の外部機器への特殊効果付き画像の送受信が行われることもある。これは、例えば、USB接続されているPC（パソコン）のインターネット接続機能を介して、インターネット上の別の機器への送受信を行うということが考えられる。

【0140】

<端末装置間のコミュニケーション>

以上説明した本実施形態における端末装置を複数用いて、これら端末装置のそれぞれにおいて、リアルタイムに取得したカラー画像に前述したように特殊効果を施して特殊効果付き画像を生成し、それを用いて、当該複数の端末装置間でコミュニケーションを行うこと  
20  
が可能となる。

【0141】

それでは、本実施形態の端末装置を用いて実現される通信システムについて、幾つかの具体例を挙げながら説明する。

【0142】

現在、カメラ機能付き携帯電話が広く普及している。これは、従来の携帯電話の機能に加えて、写真や動画を撮影して楽しむという機能、撮影物をメールなどに添付して送付すること、他人とコミュニケーションを行う機能、カメラを用いてうつした自分の姿を相手にリアルタイムに送信し、同時に相手のカメラ機能付き携帯電話でうつされた相手の姿をリアルタイムに受信することで、双方の画像を見ながらコミュニケーションを行うテレビ電話の機能などが実用化されている。  
30

【0143】

本実施形態の通信システムは、例えば、これに置き換わるかたちで実現される。上で説明したような画像コミュニケーションを行うことが可能な新たな携帯電話システムというイメージである。

【0144】

例えば、ユーザは、この新たな携帯電話システムを構成する本実施形態にかかる端末装置としての携帯電話を用いて、自分の姿や風景、気になった物などを撮影する。そして、ユーザは、その画像に対応する奥行き情報を介して生成された特殊効果付き画像（この画像には、もちろん、動画の場合もある）を得る。これを、携帯電話の待ち受け画面などに  
40  
使用し楽しむ。また、この画像（または、動画）をメールに添付して相手に送り、コミュニケーションの手段として用いる。さらに、テレビ電話といったリアルタイムな画像コミュニケーション中に、自分の感情を相手にうまく伝えたり、コミュニケーションにエンターテインメント性をもたせるために、前述同様にして特殊効果を付加したりすることができる。

【0145】

それでは、テレビ電話の例を用いて一連の流れを図33、図34に示すフローチャートを参照して具体的に説明する。いま、2人の人物A、Bが双方とも、図1に示した構成を有する携帯電話を保持している。人物A、Bともに、お互いに自分の顔を撮影している。何も特殊効果が付加されていない状態では、お互いの機器の通信部5を介して、人物AとB  
50

の携帯電話は相互に接続されており、人物Aの顔の画像は、人物Bの画像提示部4を通して提示され、逆に、人物Bの顔の画像は、人物Aに提示されている。これによって、画像を用いたテレビ電話が実現されている。そして、あるとき、人物Aは、自分の顔の周りに蝶々が飛んでいるような特殊効果を付けようと思いたち、自分の顔の画像を当該携帯電話で撮影する。これにより、画像取得部1は画像奥行き情報を含む画像（すなわち、奥行きカラー画像）を取得し（ステップS1）。特徴抽出部2では、奥行き情報などを基に当該顔画像中の3次元的な特徴を抽出する（ステップS2）。人物Aが、蝶々が飛んでいる画像を選択すると、加工部3では、当該顔画像に抽出された3次元的な特徴を基に、顔の回りを蝶々が飛び回っているように選択された画像を合成し、特殊効果付き画像を生成する（ステップS3）。この特殊効果付き画像を画像提示部4に表示するとともに（ステップS4）、人物Aの送信指示により、当該特殊効果付き画像を人物Bに送信する（ステップS5）。人物Bの所持する携帯電話では、通信部5にて当該特殊効果付き画像を受信すると（ステップS6）、それを画像提示部4に表示するので（ステップS7）、人物Bは、その特殊効果付きの人物Aの顔を見ることができ、このように、従来のテレビ電話では無かった、エンターテインメント性をコミュニケーションに付加することが可能となる。

#### 【0146】

本実施形態は、携帯電話のような場合だけではない。次に、別の例を説明する。従来から、パソコン（PC）にUSB接続のカメラを接続し、景色などを常に撮影しておく。そして、そのPCをインターネットにつなぎ、撮影しているリアルタイムの画像を一般に公開するということが行われている。これは、通常、ライブカメラ、または、定点観測カメラなどと呼ばれることが多いサービスである。撮影されているものは、自分の部屋の様子、ペットの様子、経営している店の客の入り具合の状態、渋谷の町並みの様子など多岐多様に渡っている。

#### 【0147】

このような使用方法において、上記のUSB接続カメラの代わりに、本実施形態にかかる端末装置をPCに接続することで、このような画像に特殊効果を付加することができる。従って、例えば、いま、ある商品を売っている店が、商品の画像をライブカメラを通して、インターネット上に公開しているとする。そして、カメラの向きを変えながら（パン動作）幾つかの商品を撮影しているとする。この際、いわゆる「一押し」のある商品が撮影された際に、その商品の周りに、「一押し」を表す仮想物体（オブジェクト）を図20のように回す特殊効果をかけ、新製品の商品が撮影された際には、当該新製品を表す仮想物体（オブジェクト）が回るような特殊効果をかける、といったことができる。このように、ライブカメラなどの画像に、エンターテインメント性を持たし、さらに、「一押し」、「新製品」などといった付加情報を提供することができる。

#### 【0148】

（第1の実施形態の第1の変形例）

図25は、第1の変形例に係る端末装置の要部の構成例を示したものである。

#### 【0149】

図25に示した端末装置と図1に示した端末装置との違いは、図25では、特徴抽出部2が、画像取得部1で取得した奥行きカラー画像に含まれる奥行き情報を基に撮影対象の3次元的な特徴を前述同様にして抽出するとともに、さらに、通信部5で受信した奥行きカラー画像からも、そこに含まれる奥行き情報を基に撮影対象の3次元的な特徴を抽出するものである。

#### 【0150】

また、図25の加工部3は、特徴抽出部2で抽出された特徴をもとに、画像取得部1や通信部5で受信した奥行きカラー画像（あるいはカラー画像）や奥行き情報を含む特殊効果付き画像に、前述同様にして特殊効果を施すように構成されている。

#### 【0151】

ここで、特殊効果付き画像は、もともと奥行き情報の対応付けられているカラー画像や奥行きカラー画像から生成されているので、特殊効果付き画像にも奥行き情報は対応付けら

10

20

30

40

50

れている（含まれている）。

#### 【0152】

このような構成にすることで、画像取得部1で得られた奥行きカラー画像だけでなく、通信部5で受信した奥行きカラー画像や特殊効果付き画像に関しても特殊効果を施すことが可能となる。

#### 【0153】

次に、図25に示した構成を有する端末装置の動作について、図35に示すフローチャートを参照しながら、先の携帯電話の例を用いて具体的に説明する。

#### 【0154】

いま、2人の人物が双方とも、図25に示したような構成の（第1の変形例に係る機能を持った）携帯電話を保持しており、2人の間で、テレビ電話を用いたコミュニケーションを行っているとする。この際、第1の実施形態では、自分の撮影している画像に対して特殊効果をかけることが可能であったが、第1の変形例では、これに加え、通信部5では、相手から送られてくる奥行き情報を含む画像を受信し（ステップ811）、これを表示するとともに（ステップ812）、当該受信した画像の奥行き情報などを基に、当該受信した画像の3次元的な特徴を抽出する（ステップ813）。従って、受信側の加工部3では、当該受信した画像に対して特殊効果を付加することが可能となる（ステップ814）。所望の特殊効果を付加して生成された特殊効果付き画像は、画像提示部4に提示するとともに、さらに、再び相手に送信することもできる（ステップ816）。

#### 【0155】

従って、例えば、双方とも、自分の顔をうつしてテレビ電話を用いたコミュニケーションを行っている最中に、相手の顔にいたずらをするができる。例えば、よくテレビのコント番組では、怒りの表現や罰ゲームとして、人物の上に、「たらい」を落として笑いをとることがある。これと同じように、相手とのコミュニケーションの中で、なにか不快な思いをした際に、相手の顔画像に対して、「たらい」を落とすような特殊効果を施し、自分が見て楽しむ、さらに、それを相手におくって、楽しませながら、注意を促す、といった新たな画像コミュニケーションが可能となる。

#### 【0156】

（第1の実施形態の第2の変形例）

図26は、第2の変形例に係る端末装置の要部の構成例を示したものである。

#### 【0157】

図26に示した端末装置と図1に示した端末装置との違いは、図26では、図1に示した構成に、さらに、仮想物体などの複数の付加画像データ（例えば、CGデータや3DCGデータ）を記憶する付加画像記憶部7と、この中から所望の付加画像や特殊効果を選択するための特殊効果選択部8とが追加された構成になっている。

#### 【0158】

特殊効果選択部8は、付加画像記憶部7に記憶されている付加画像を参照し、付加可能な付加画像や特殊効果の種類をユーザに呈示する。ユーザは、呈示されたものの中から付加したい付加画像あるいは特殊効果を選択し、この情報が加工部3に渡される。加工部3は、ユーザが選択した付加画像や特殊効果の種類に従った特殊効果を付加する。また、特殊効果選択部8では、ユーザに付加画像や特殊効果の種類を選択させるのではなく、ランダムに付加画像や特殊効果を選択することもある。

#### 【0159】

このように変形することで、ユーザは、取得したカラー画像に合成する仮想物体や、特殊効果を選択して、ユーザの好みに応じた特殊効果付きカラー画像を生成することが可能となる。

#### 【0160】

（第1の実施形態の第3の変形例）

図27は、第3の変形例に係る端末装置の要部の構成例を示したものである。第3の変形例は、上記第2の変形例をさらに変形したものであり、図27に示した端末装置と図26

10

20

30

40

50

に示した端末装置との違いは、図 27 では、図 26 に示した構成に、さらに、画像取得部 1 で取得した奥行きカラー画像（あるいはカラー画像）に映っているシーンの特徴を解析するシーン解析部 8 が追加された構成になっている。

#### 【0161】

シーン解析部 8 は、画像取得部 1 で取得した奥行きカラー画像（あるいはカラー画像）に映っているシーンの特徴を解析するものである。シーンの特徴とは、シーンに映っているものの特定、シーンの色的な雰囲気、シーンの構図、などを指す。例えば、シーンに映っている物体が人物であるならば、人物であると判別する。そして、その情報を特殊効果選択部 6 に与える。特殊効果選択部 6 では、特殊効果の種類を人物に付加することがふさわしいものに限定する。また、シーンに映っているものが縦長の場合は、その周りを横に回転する仮想物体（オブジェクト）の画像を付加（例えば合成）すると効果があるが、非常に、横長の場合は、その周りを縦に回転する仮想物体（オブジェクト）の画像を付加した方がよい場合もあるし、回転する動きのある仮想物体でない別の仮想物体の画像を合成するなどの特殊効果を選択した方が効果的であることもある。また、シーンの全体的な色合いから、仮想物体などの付加画像の形状や色を様々に変えた方がよいこともある。例えば、全体的に赤いものが多く映っているシーンに、赤い仮想物体を合成しても効果が少ないであろう。さらに、シーンにうつっている物体の配置などの画像（シーン）自体の構図に応じて、仮想物体の動きを変化させるようにしてもよい。

#### 【0162】

以上のように、図 27 に示した構成により、シーン解析部 8 にて、シーンの特徴を解析することと、当該シーンの特徴に合わせて、付加する仮想物体などの動き、色、形状といったパラメータなどを変化させたり、当該シーンの特徴に対応する特殊効果の種類を選択したりして、より効果の高い特殊効果を付加することができるのである。

#### 【0163】

（第 1 の実施形態の第 4 の変形例）

第 4 の変形例に係る端末装置は、図 1 に示した端末装置の構成、図 25 に示した第 1 の変形例の構成、図 26 に示した第 2 の変形例の構成、図 27 に示した第 3 の変形例の構成のいずれかに、さらに、通信部 5 で通信する画像を選択することが可能な選択部 9 が追加された構成となっている。

#### 【0164】

図 28 は、第 4 の変形例に係る端末装置の要部の構成例を示したものであり、図 28 では、図 25 に示した第 1 の変形例の構成に選択部 9 を追加して構成される端末装置を示している。

#### 【0165】

図 28 に示したように、選択部 9 を追加することにより、通信部 5 を介して他の端末装置と交換するデータを選択することが可能となる。

#### 【0166】

第 1 の変形例のところで用いた例を用いて、第 4 の変形例によって得られる効果について説明する。第 1 の変形例と同様の例を考える。いま、2 人の人物が双方とも第 4 の変形例に係る図 28 に示したような構成の機能を持った携帯電話を保持しており、2 人の間で、テレビ電話を用いたコミュニケーションを行っているとする。そして、双方とも、自分の顔をうつしてテレビ電話を用いたコミュニケーションを行っている。この最中に、様々な特殊効果を用いたコミュニケーションを行うが、相手の顔にいたずらをする特殊効果のみは、相手に送信しないなどといった選択を行う。これにより、普段は、特殊効果付き画像を常に相手に送っているが、なにか不快な思いをした際に、相手の顔画像に対して、「たらい」を落とすような特殊効果を施し、その特殊効果付き画像のみは、自分が見て楽しむだけで、相手には送らない、といったことが可能となる。

#### 【0167】

これは、1 対 1 のコミュニケーションだけでなく、多人数とのコミュニケーション（1 対多数、多数対多数）において、さらに効果がある。例えば、ある人物 D が、同時に 3 人

10

20

30

40

50

の人物 A、B、C と、図 28 に示した構成を有するテレビ電話器を用いたコミュニケーションを行っていたとする。あるとき、人物 D は、人物 A の言動に対して不快な思いをして、「たらい」を落とす特殊効果を施す選択をした。しかし、A にそれを送ったのではけんかになると考えたため、人物 B、C のみにそれを送り、自分の感情を伝えた、などといった送信対象の選択が可能となる。

【0168】

以上では、送信に対する選択の説明をしたが、受信に対しても、同様に選択し、特定の人からの特殊効果付き画像を受信しないようにしたり、特定の特殊効果の施された特殊効果付き画像は受信しないようにしたりするようにしてもよい。

【0169】

(第 1 の実施形態の第 5 の変形例)

特殊効果付きカラー画像には、当該画像に適した効果音を付加することも可能である。

【0170】

特殊効果付きカラー画像には、当該画像に適した効果音を付加するための効果音付加部を、図 1 に示した構成、図 25 に示した第 1 の変形例の構成、図 26 に示した第 2 の変形例の構成、図 27 に示した第 3 の変形例の構成のいずれかにさらに追加することもできる。

【0171】

これにより、特殊効果付きカラー画像に効果音をつけることができ、特殊効果をさらに高めることができる。例えば、前述の第 1 の実施形態では、仮想物体が人物に当たって跳ね返り、その際に衝突したことを表す「星」など仮想物体の画像を合成する特殊効果の例を説明したが、図 1 に示した構成において、上記効果音付加部（図示せず）を追加すると、この衝突の際に、衝突音を付加することができ、当該特殊効果の効果をあげることができる。

【0172】

以上説明したように、上記第 1 の実施形態によれば、撮影対象の画像（例えば、ここでは自然光画像）とともに、（各画素に奥行き情報を含む反射光画像を取得することにより）当該自然光画像の奥行き方向の情報（奥行き情報）を取得することができるので、この奥行き情報を用いて、自然光画像中の撮影対象の 3 次元的な形状や撮影対象の 3 次元的な位置関係などの 3 次元的な特徴を抽出し、この 3 次元的な特徴を基に、衝撃などの効果表現や仮想物体を表現した画像（付加画像）を合成するなどして、画像中の撮影対象の 3 次元的な特徴を生かした特殊効果を付加することができる。

【0173】

(第 2 の実施形態)

次に、本発明の第 2 の実施形態について説明する。

【0174】

<全体の構成>

図 29 は、第 2 の実施形態に係る通信システムの全体の概略構成を示したものである。図 29 に示したように、この通信システムは、無線通信システムであって、例えば、図 1 に示した構成を有する複数の端末装置（例えば、ここでは、第 1 の端末 201 と第 2 の端末 202 の 2 つ）と、第 1 の端末 201 と無線接続する基地局装置（BS）211 と第 2 の端末 202 と無線接続する基地局装置（BS）212 と、基地局装置 211、212 とサーバ装置 200 を接続する所定の通信網（ネットワーク）とから構成されている。サーバ装置 200 は、基地局 211、212、ネットワークを介して第 1 の端末 201 と第 2 の端末 202 と互いに通信可能に接続されている。

【0175】

第 1 の端末 201 と第 2 の端末 202 とは、前述したように、特殊効果付き画像や特殊効果のついていない画像を互いに送受信することによって画像コミュニケーションが行える。なお、第 1 の端末 201 と第 2 の端末 202 はそれぞれ図 1 に示した構成を有しているため、それぞれの端末ではカラー画像の他に、当該カラー画像に対応する奥行き情報を、すなわち、ここでは、奥行き情報を含むカラー画像としての奥行きカラー画像を取得することが

10

20

30

40

50



できる。

【0176】

第1の端末201と第2の端末202との間で、画像コミュニケーションを行う際には、各端末の通信部5からは、特殊効果付き画像あるいは特殊効果のついていない画像とともに、当該画像に対応する実行き情報も送信する。なお、ここでは、前述同様、送信する画像自体に実行き情報が含まれているものとする。

【0177】

また、第1の端末201と第2の端末202との間で、特殊効果付き画像を送受信する場合には、各端末の通信部5からは、当該実行き情報を含む特殊効果付き画像とともに、各端末の特徴抽出部2で抽出された3次元的な特徴も送信される。

10

【0178】

また、サーバ装置200は、第1の端末201と第2の端末202との間で、画像コミュニケーションが行えるサービスを提供するアプリケーションにより設置および運営されており、第1の端末201と第2の端末202との間で当該画像コミュニケーションを行う際には、第1の端末201と第2の端末202から送信される画像や、実行き情報などは、必ず、サーバ装置200により受信されて、このサーバ装置200を経由して相手端末へ送信されるようになっている。

【0179】

さて、この第2の実施形態では、上記サーバ装置200において、第1の端末201と第2の端末202から送信される画像に、当該画像に対応する実行き情報を基に、付加画像すなわち、ここでは、広告画像を合成する点に特徴がある。

20

【0180】

図30は、サーバ装置200の構成例を示したもので、受信部10と特徴抽出部2と加工部3と送信部11とから構成されている。なお、図30において、図1と同一部分には同一符号を付し、異なる部分について説明する。すなわち、図30に示すサーバ装置200は、図1の端末装置の画像取得部1と画像提示部4と通信部5のそれぞれに代えて、第1の端末201や第2の端末202から送信されてきた画像（実行き情報を含む、特殊効果付き画像あるいは特殊効果のついていない画像）を受信する受信部10と、サーバ装置200の加工部3で、受信部10で受信した画像に広告画像の合成などの特殊効果を施すことにより生成された、広告付き画像を当該受信した画像の宛先である第1の端末201あるいは第2の端末202へ送信する送信部11を具備している。

30

【0181】

次に、図36に示すフローチャートを参照して、サーバ装置200の処理動作について説明する。

【0182】

サーバ装置200の受信部10は、第1の端末201や第2の端末202から送信されてきた実行き情報を含む画像（特殊効果付き画像あるいは特殊効果のついていない画像）を受信すると（ステップS21）、サーバ装置200の特徴抽出部2は、図1の端末装置の特徴抽出部2と同様にして、受信部10で受信した画像と当該画像に対応する実行き情報とから、当該画像中の3次元的な特徴を抽出する（ステップS22）。

40

【0183】

サーバ装置200の加工部3では、受信部10で受信した画像と、当該抽出された3次元的な特徴とを基に、広告画像を合成するなどして、当該受信した画像に特殊効果を施す（ステップS23）。この加工部3で広告画像を合成するなどして生成された広告付き画像は、送信部11から、当該受信した画像のもとともに宛先である第1の端末201あるいは第2の端末202へ送信される（ステップS24）。

【0184】

ここで、サーバ装置200の加工部3について説明する。加工部3は、受信部10で受信した画像（既に第1、第2の端末201、202にて特殊効果のつけられている特殊効果付き画像、あるいは、特殊効果のつけられていない一般的な画像）から特徴抽出部2によ

50

って得られた3次元的な特徴を考慮して、受信部10で受信した当該画像に任意の広告画像をさらに合成するなどして特殊効果を施すものである。

【0185】

具体的には、当該画像に、さらに、CG（コンピュータグラフィックス）などで表現された広告を合成することで、広告付きの画像を生成する。この際、特徴抽出部2で抽出された、画像中の撮影対象（物体）の凹凸といった3次元的な形状、シーン内にどのような物体が存在するか（どのような領域に分割できるか）、シーンにおける各物体の位置関係、それぞれの物体の立体形状、などといった3次元的な特徴を基に、図1の加工部3と同様に、広告画像と受信した画像中の物体の前後関係などの位置関係や衝突などを判別し、広告を合成する。

10

【0186】

ここで、図7に示した画像が受信部10で受信されたとし、この図7に示した画像（以下、受信画像と呼ぶ）に広告（画像）を合成する場合を例にとり、サーバ装置200の加工部3の処理動作について説明する。

【0187】

最も簡単な広告の付加方法は、受信画像の前景に広告が架からないように、背景部分に広告を合成するものである。図31は、図7の受信画像に広告画像を合成した様子を示したものである。図14に示したように、特徴抽出部2で抽出された3次元的な特徴によれば、容易に前景と背景を判別することができるのは前述したとおりである。この3次元的な情報をを用いて、広告を付加する訳である。こうすることで、前景である人物が動いても、人物には広告が架からないように、常に背景に広告を提示しておくことが可能となる。こうすることで、コミュニケーションに悪影響を及ぼすことなく、常に広告を提示しておくことが可能となる。

20

【0188】

さらに、受信画像中の端末装置の加工部3で付加された（合成された）仮想物体に広告を追加しても効果的である。例えば、受信画像に、図17で説明した人物の後ろを仮想物体「球」が動くという特殊効果が既に施されていたとする。サーバ装置200の加工部3では、この仮想物体「球」自体に広告を貼り付ける、あるいは、図32に示すように、「球」に広告をぶら下がるように、広告を合成する。こうすることで、「球」の動きにあわせて、広告自体も画面内を動き回ることになる。つまり、受信画像に含まれる端末側で付加された特殊効果によるエンターテインメント効果の中に広告効果も含めてしまうわけである。これにより、ユーザに悪影響を与えることなく、エンターテインメント効果を与える中でさりげなく広告提示をすることが可能となる。

30

【0189】

また、仮想物体自体が広告となっていてよく、サーバ装置200の加工部3では、広告と一体化した仮想物体を受信画像に合成する。広告と一体化した仮想物体とは、例えば、新発売の商品のCG画像などのようなものが挙げられる。

【0190】

以上説明した広告の付加方法はあくまでも一例であり、これに限定されるものではない。サーバ装置200の特徴抽出部2で抽出された3次元的な特徴を用いて、自由に広告を付加することができる。例えば、図22には、頭にペンキが落ちてくる特殊効果の例を示したが、これが、ペンキではなく、ある商品のパッケージが落ちてきてもよいし、商品名や発売日などを示した文字が落ちてきてもよい。図18に、仮想物体と衝突判定をし、その際に星などの仮想物体をつける例を示したが、この星のなかに広告が含まれて提示されてもよい。

40

【0191】

また、画面内に2つ以上の別の広告を出すことも可能であるし、受信画像が動画の場合には、シーンによって広告の内容が変わってもよい。例えば、シーンに写っている人物が女性ならば、女性向けの広告を、男性ならば男性向けの広告を付加するといったようにである。

50

## 【0192】

さらに、第1の端末201や第2の端末202では、もともと、それぞれの特徴抽出部2にて3次元的な特徴が抽出されているわけであるから、この抽出された3次元的な特徴情報も、画像（実行き情報を含む、特殊効果付き画像あるいは特殊効果のついていない画像）とともに、必ず送信するようにすれば、サーバ装置200には特徴抽出部2は持つ必要はなくなる。そして、サーバ装置200では、図37に示すように、受信部10で受信した画像（受信画像）に対し、同じく受信部10で受信した当該受信画像に対応する特徴情報を基に前述同様にして広告画像を合成することができ（ステップ831～ステップ832）。

## 【0193】

第2の実施形態に係るサーバ装置200を用いた通信システムによれば、第1の実施形態で説明したような画像コミュニケーションに加え、送信する画像中に広告を効果的に追加した画像を提供することが可能となる。これにより、広告をユーザ間のコミュニケーションの邪魔にならない位置に提示したり、広告を既にある仮想物体中に組み込むことで、広告にエンターテインメントを持たせるとともに、ユーザに悪影響を与えずに広告を提示することが可能となる。

## 【0194】

なお、以上の各実施形態やその変形例は、適宜組み合わせることで実施することが可能である。

## 【0195】

本願発明の実施形態における処理をコンピュータで実行可能なプログラムで実現し、このプログラムをコンピュータで読みとり可能な記憶媒体として実現することも可能である。例えば、図1の端末装置の構成のうち、画像取得部1の一部（例えば、自然光画像撮像部103と反射光画像撮像部108と撮像動作制御部104）は、前述したように、1つの集積回路に収められていることが望ましいので、少なくとも、それ以外の画像取得部1と特徴抽出部2と加工部3はソフトウェアとして実現することも可能である。また、上記実施形態に係る端末装置の図33～図35に示した処理動作を行わせるためのプログラムを当該端末装置にインストールすることにより、実現することができる。

## 【0196】

なお、記憶媒体としては、磁気ディスク、フロッピーディスク、ハードディスク、光ディスク（CD-ROM、CD-R、DVD等）、光磁気ディスク（MO等）、半導体メモリ等、プログラムを記憶でき、かつコンピュータまたは組み込みシステムが読みとり可能な記憶媒体であれば、その記憶形式は何れの形態であってもよい。

## 【0197】

また、記憶媒体からコンピュータや組み込みシステムにインストールされたプログラムの指示に基づきコンピュータ上で稼働しているOS（オペレーションシステム）や、データベース管理ソフト、ネットワーク等のMW（ミドルウェア）等が本実施形態を実現するための各処理の一部を実行してもよい。

## 【0198】

さらに、上記記憶媒体は、コンピュータあるいは組み込みシステムと独立した媒体に限らず、LANやインターネット等により伝達されたプログラムをダウンロードして記憶または一時記憶した記憶媒体も含まれる。

## 【0199】

また、記憶媒体は1つに限られず、複数の媒体から本実施形態における処理が実行される場合も、本発明における記憶媒体に含まれ、媒体の構成は何れの構成であってもよい。

## 【0200】

また、本願発明におけるコンピュータまたは組み込みシステムは、記憶媒体に記憶されたプログラムに基づき、本実施形態における各処理を実行するためのものであって、パソコン、マイコン等の1つからなる装置、複数の装置がネットワーク接続されたシステム等の何れの構成であってもよい。

## 【0201】

10

20

30

40

50

また、本願発明におけるコンピュータとは、パソコンに限らず、情報処理機器に含まれる演算処理装置、マイコン等も含み、プログラムによって本願発明の機能を実現することが可能な機器、装置を総称している。

#### 【0202】

また、本発明は、上記第1～第2の実施形態に限定されるものではなく、実施段階ではその要旨を逸脱しない範囲で種々に変形することが可能である。さらに、上記実施形態には種々の段階の発明が含まれており、開示される複数の構成要件における適宜な組み合わせにより、種々の発明が抽出され得る。例えば、実施形態に示される全構成要件からいくつかの構成要件が削除されても、発明が解決しようとする課題の欄で述べた課題（の少なくとも1つ）が解決でき、発明の効果の欄で述べられている効果（の少なくとも1つ）が得られる場合には、この構成要件が削除された構成が発明として抽出され得る。

10

#### 【0203】

##### 【発明の効果】

以上説明したように、本発明によれば、撮影した画像（動画像、静止画像を含む）中の複数の撮影対象のそれぞれの3次元的な形状や位置関係などの3次元的な特徴を基に、広告画像や所望の仮想物体の画像を合成するなどの加工が行える。

##### 【図面の簡単な説明】

【図1】本発明の第1の実施形態に係る端末装置の要部の構成例を概略的に示す図。

【図2】通常の白黒画像と、奥行き白黒画像を説明するための図。

【図3】画像取得部の構成例を概略的に示す図。

20

【図4】反射光画像撮像部の構成例を示した図。

【図5】人間の手を撮影対象として、反射光画像撮像部で撮影された反射光画像に含まれる奥行き情報から得られる、当該反射光画像中の手の3次元的なイメージを表した図。

【図6】画像取得部で取得された反射光画像（（a）図）と自然光画像（（b）図）の一例を示した図。

【図7】特徴抽出部での処理対象である奥行きカラー画像の一例を白黒画像で示した図。

【図8】図7の奥行きカラー画像中の物体の3次元的な位置関係を概略的に示す図。

【図9】図7の奥行きカラー画像中の背景部分を示した図。

【図10】図7の奥行きカラー画像中の人物部分を示した図。

【図11】図7の奥行きカラー画像中の腕部分を示した図。

30

【図12】図7の奥行きカラー画像中の缶部分を示した図。

【図13】図7の奥行きカラー画像中の物体の奥行き方向の一関係を示した図。

【図14】図7の奥行きカラー画像中の奥行き方向の位置関係と前景・背景を示した図。

【図15】図7の奥行きカラー画像中の奥行き方向の不連続点（ジャンプエッジ）から当該画像中の物体の詳細な位置関係を求める方法を説明するための図。

【図16】図7の奥行きカラー画像中の各物体の奥行き方向の位置を特定する方法を説明するための図。

【図17】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像の一例を説明するための図。

【図18】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像の他の例を説明するための図。

40

【図19】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像のさらに他の例を説明するための図。

【図20】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像のさらに他の例を説明するための図。

【図21】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像のさらに他の例を説明するための図。

【図22】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像のさらに他の例を説明するための図。

【図23】図7の奥行きカラー画像に特殊効果を加えた結果を示した特殊効果付き画像の

50

さらに他の例を説明するための図。

【図24】画面提示部に、画像を表示するための画面分割の例を示した図。

【図25】本発明の第1の実施形態の第1の変形例に係る端末装置の要部の構成例を概略的に示す図。

【図26】本発明の第1の実施形態の第2の変形例に係る端末装置の要部の構成例を概略的に示す図。

【図27】本発明の第1の実施形態の第3の変形例に係る端末装置の要部の構成例を概略的に示す図。

【図28】本発明の第1の実施形態の第4の変形例に係る端末装置の要部の構成例を概略的に示す図。

【図29】本発明の第2の実施形態に係る通信システムの全体の構成を概略的に示した図

10

【図30】サーバ装置の要部の構成例を概略的に示した図。

【図31】広告付き画像の一例を説明するための図。

【図32】広告付き画像の他の例を説明するための図。

【図33】端末装置の処理動作を説明するためのフローチャート。

【図34】端末装置の処理動作を説明するためのフローチャート。

【図35】端末装置の処理動作を説明するためのフローチャート。

【図36】サーバ装置の処理動作を説明するためのフローチャート。

【図37】サーバ装置の処理動作を説明するためのフローチャート。

20

【符号の説明】

1 画像取得部

2 特徴抽出部

3 加工部

4 画像提示部

5 通信部

6 特殊効果選択部

7 付加画像記憶部

8 シーン解析部

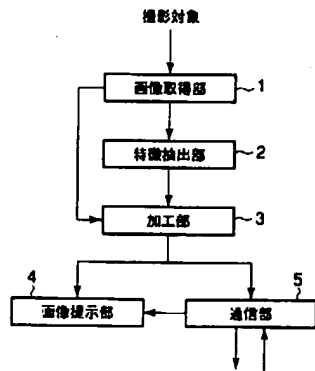
9 選択部

30

10 受信部

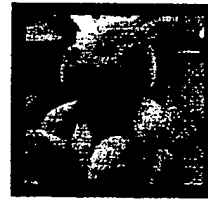
11 送信部

【図 1】

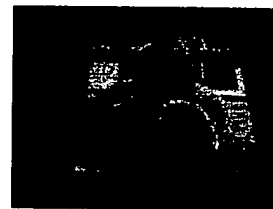


【図 2】

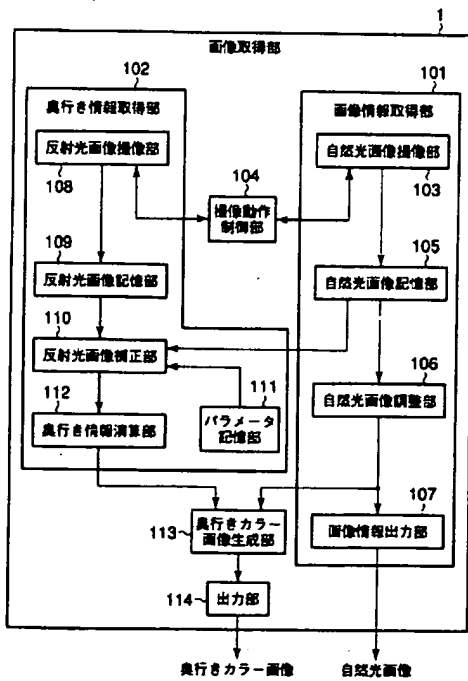
(a)



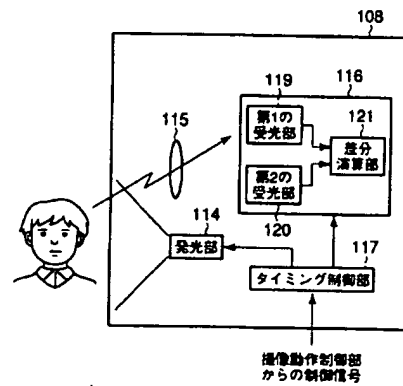
(b)



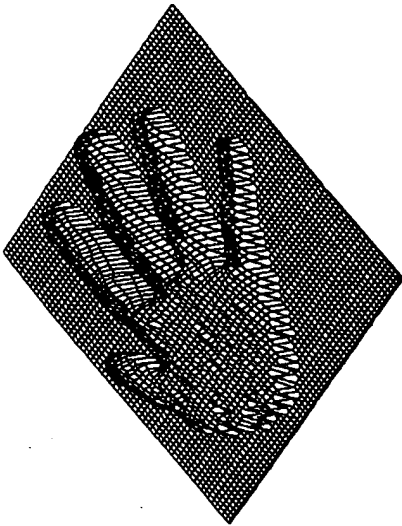
【図 3】



【図 4】



【図 5】



【図 6】

(a)

0	0	0	0	0
0	0	0	0	0
0	0	255	0	0
0	0	250	0	0
0	0	250	0	0
0	220	250	0	220
0	220	220	0	220
0	220	220	200	200

P2(5, 8)

(b)

(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)	(0, 0, 0)
(0, 0, 0)	(0, 0, 0)	(r2, g2, b2)	(0, 0, 0)	(0, 0, 0)
(0, 0, 0)	(0, 0, 0)	(r2, g2, b2)	(0, 0, 0)	(0, 0, 0)
(0, 0, 0)	(0, 0, 0)	(r2, g2, b2)	(0, 0, 0)	(0, 0, 0)
(0, 0, 0)	(r1, g1, b1)	(r2, g2, b2)	(0, 0, 0)	(r2, g2, b2)
(0, 0, 0)	(r1, g1, b1)	(r2, g2, b2)	(0, 0, 0)	(r2, g2, b2)
(0, 0, 0)	(r1, g1, b1)	(r2, g2, b2)	(r3, g3, b3)	(r3, g3, b3)

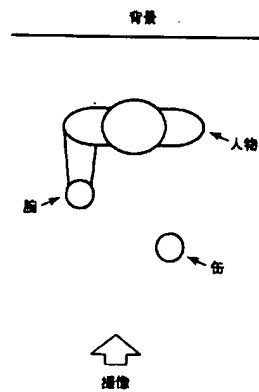
(R, G, B)

P1(5, 8)

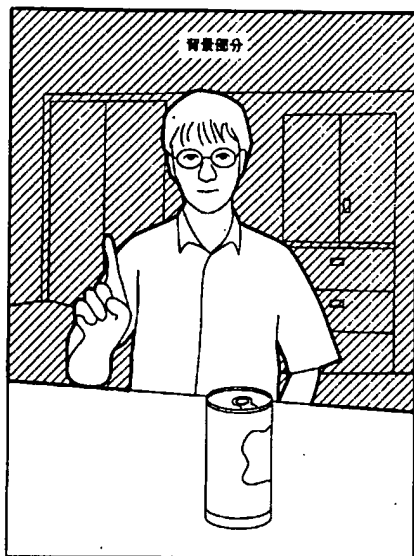
【図 7】



【図 8】



【図 9】



【図 10】



【図 11】

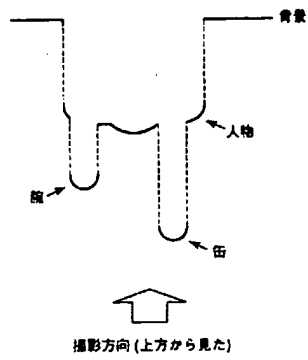


【図 12】

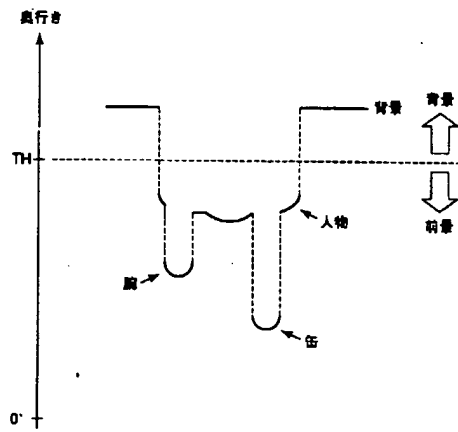




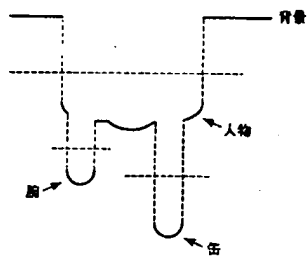
【図 13】



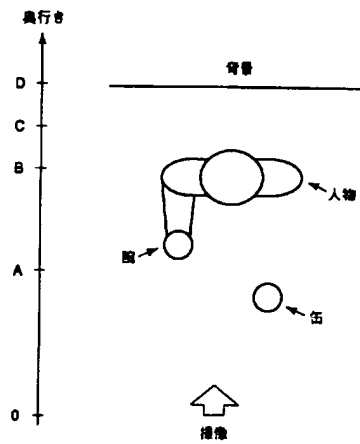
【図 14】



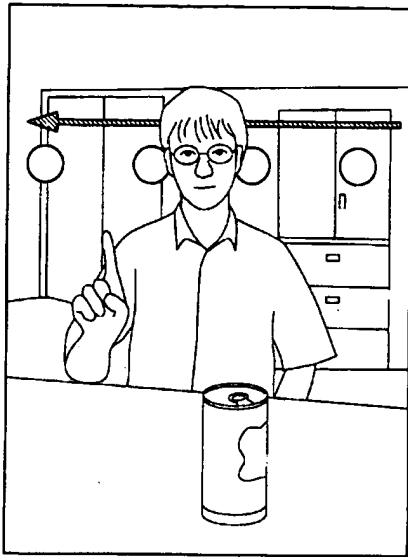
【図 15】



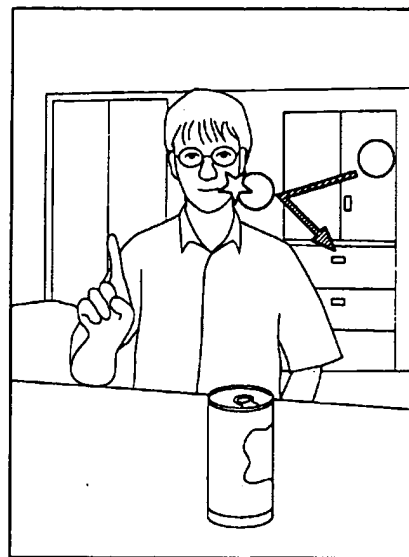
【図 16】



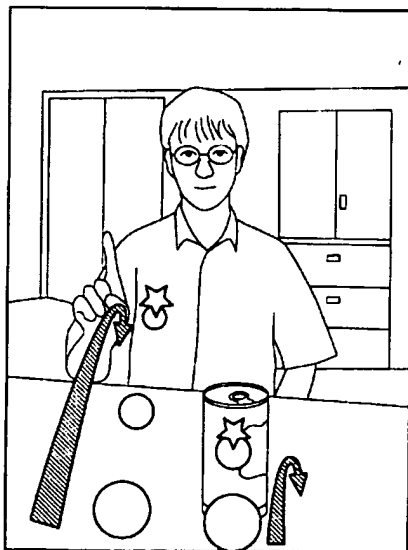
【図 17】



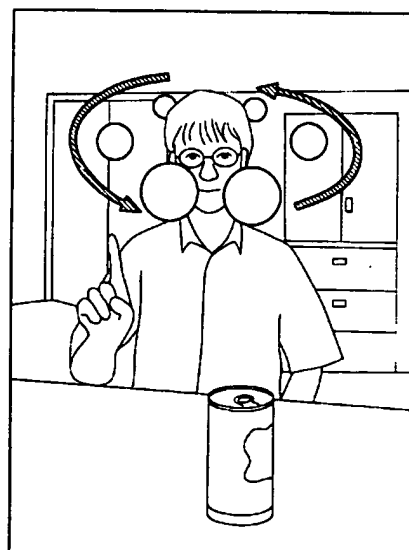
【図 18】



【図 19】



【図 20】



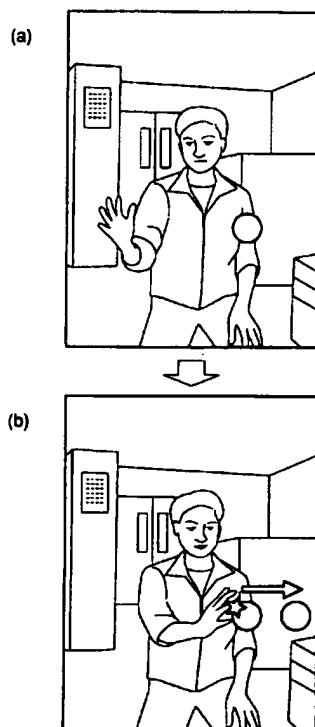
【図 21】



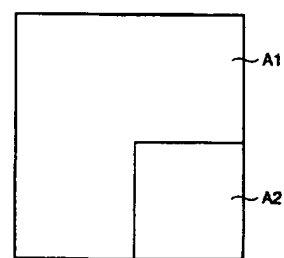
【図 22】



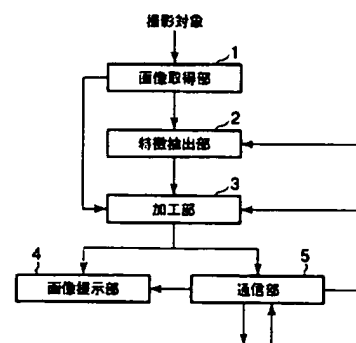
【図 23】



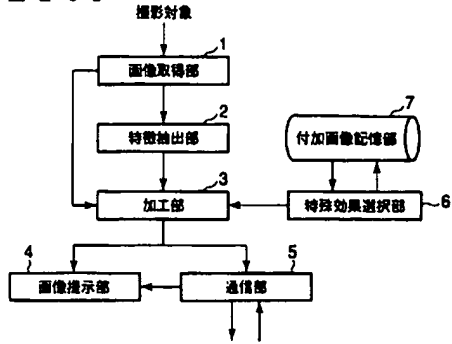
【図 24】



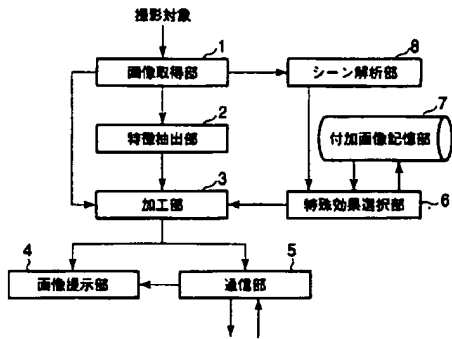
【図 25】



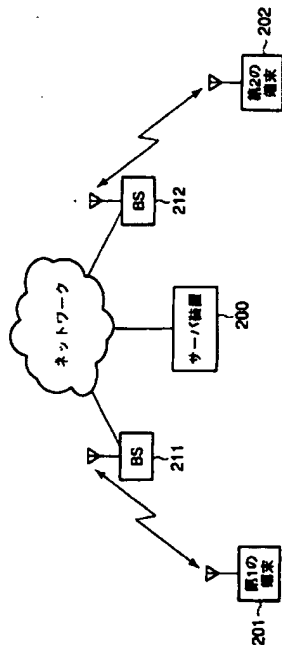
【図26】



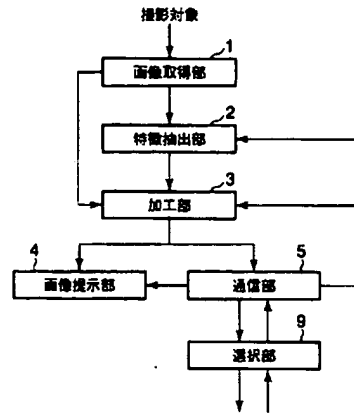
【図27】



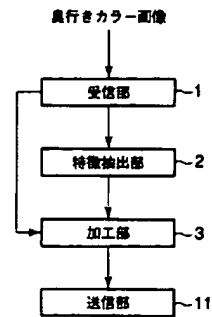
【図29】



【図28】



【図30】



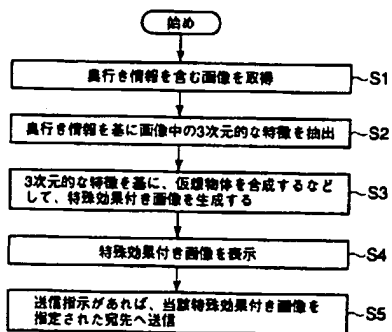
【図 3 1】



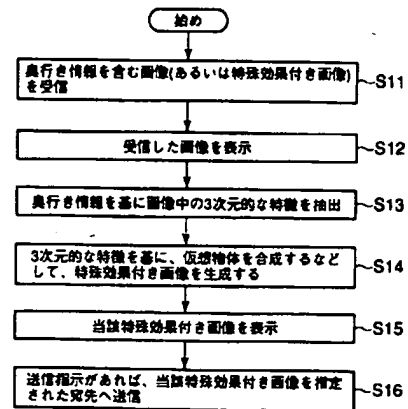
【図 3 2】



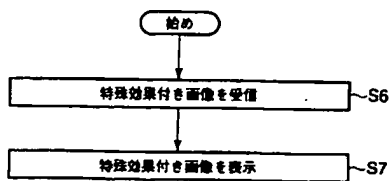
【図 3 3】



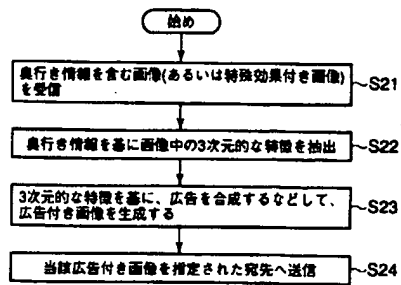
【図 3 5】



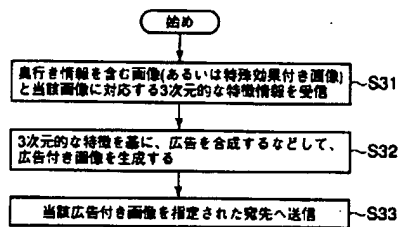
【図 3 4】



【図 36】



【図 37】



フロントページの続き

(51)Int. Cl.

F I

テーマコード (参考)

G 0 6 T 7/60 1 8 0 B

G 0 6 T 15/70 A

(74)代理人 100070437

弁理士 河井 将次

(72)発明者 三原 功雄

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 沼崎 俊一

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 原島 高広

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 岸川 晋久

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

(72)発明者 土井 美和子

神奈川県川崎市幸区小向東芝町1番地 株式会社東芝研究開発センター内

Fターム(参考) 5B050 BA04 BA06 BA07 BA08 BA12 DA04 DA07 EA04 EA06 EA07

EA13 EA17 EA19 EA24 FA02 FA05

5B057 CA01 CA08 CA13 CA16 CB01 CB08 CB12 CB16 CE08 CE16

CF05 DA07 DA08 DA16 DB03 DB06 DB09 DC02 DC09 DC16

DC30 DC32

5L096 AA02 AA06 AA09 CA04 FA02 FA06 FA66 FA69 FA76 GA08

GA10 GA40 HA07